PHYSICS
FOR ENGINEERS
Part-2

DR. GIAS UDDIN AHMAD

HAFIZ BOOK CENTRE

Gias Uddin Ahmad

CONTENTS

CHAPTER – I

COULOMB'S LAW & ELECTRIC FIELD

1.1 Concept of charge

Let a glass rod, rubbed with silk, be hung from a long silk thread as in Fig. 1.1. If another glass rod, again rubbed with silk, is held near the rubbed end of the first rod, the rods will *repel* each other. On the other hand, a hard-rubber rod, rubbed with fur, will *attract* the glass rod. But two rubber rods rubbed with fur will repel each other. This phenomenon serves us to introduce the concept of charge. We say that as a result of rubbing, the rods acquire a new property

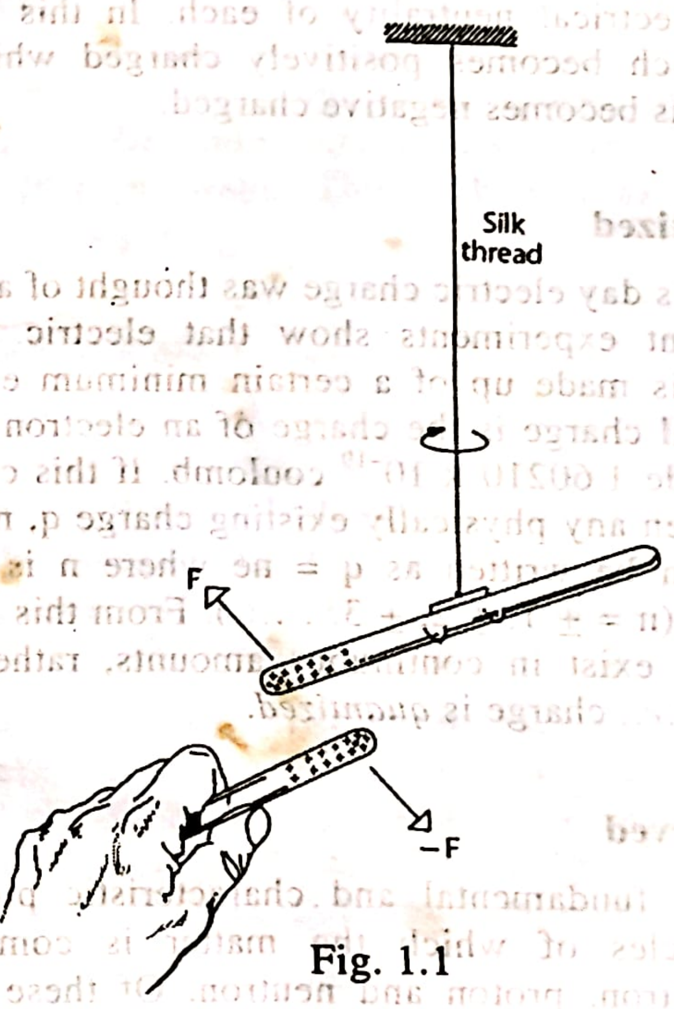


Fig. 1.1

– they become electrically charged and the charges on the two rods exert forces on one another. Clearly the charges on the glass the hard rubber rods must be different in nature. Benjamin Franklin who, among many other accomplishments, is regarded as the first

American physicist, named the kind of electric charge that appears on the glass rod as *positive* and the kind that appears on hard-rubber *negative*. The positive and negative labels and signs for electric charge were chosen arbitrarily by Franklin – but the names have remained to this day. Thus the simple experiments described above can be summed up by saying that

like charges repel each other and unlike charges attract each other.

Charge itself, however, is not created during the process of rubbing. In the light of modern view of bulk matter, an object in its normal state is electrically neutral *i.e.*, it contains equal amounts of positive and negative charge. If the two bodies like glass and silk are rubbed together, electrons are transferred from one to the other, upsetting the electrical neutrality of each. In this case electrons leave glass which becomes positively charged while silk which receives electrons becomes negative charged.

Charge is quantized

In Franklin's day electric charge was thought of as a continuous fluid. Subsequent experiments show that electric charge is not continuous but is made up of a certain minimum electric charge. This fundamental charge is the charge of an electron or proton and has the magnitude 1.60210×10^{-19} coulomb. If this charge is given the symbol e, then any physically existing charge q, no matter what its origin is, can be written as $q = ne$ where n is a positive or negative integer ($n = \pm 1, \pm 2, \pm 3, ...$). From this it follows that charge does not exist in continuous amounts, rather it exists in discrete *packets i.e.*, charge is *quantized.*

Charge is conserved

Charge is a fundamental and characteristic property of the elementary particles of which the matter is composed. These particles are electron, proton and neutron. Of these electrons are negatively charged, protons positively charged and neutrons are neutral. Although the mass of a proton is 1840 times heavier than the mass of an electron, the magnitude of its charge is the same as that of electron. The masses of proton and neutron are approximately equal.

There are then two kinds of charge, positive and negative; and an ordinary piece of matter contains equal amount of each kind. Thus ordinary matter is electrically neutral. When we say that an object is charged, we mean that either it has an excess of electrons (in which case the object is negatively charged) or an excess of protons (in which case the object is positively charged). The charge of a body, then, refers to net charge or excess charge. It is the algebraic sum of the charges present in the body and may be positive, negative or zero.

Like many other physical entities charge cannot be created or destroyed. When a glass rod is rubbed with silk, a positive charge appears on the rod. Measurement shows that a negative charge of equal magnitude appears on the silk. This suggests that charge is not created by rubbing but is merely transferred from one object to another, slightly disturbing the electrical neutrality of each object in the process. Thus the total charge of an isolated (or closed) system cannot change – the individual charges can be combined or regrouped in different ways. This is known as the *principle of conservation of charges.*

Conductors and insulators

A metal rod held in hand and rubbed with fur will not seem to develop a charge. If the same rod is provided with a glass or hard rubber handle and then rubbed with fur without touching it with hands, it will be charged. The explanation is that metals, the human body and the earth are *conductors* of electricity while glass, hard rubber, plastics, etc. are *insulators* (also called *dielectrics*).

In conductors electric charges are free to move through the material, whereas in insulators they are not. Although there are no perfect insulators, the insulating ability of fused quartz is 10^{25} times as great as that of copper; so that for many practical purposes fused quartz as well as many other materials behave as if they were perfect insulators.

There is yet another class of materials called *semiconductors* whose ability to conduct electricity is intermediate between conductors and insulators. Among the elements, silicon and germanium are well known examples. The electrical conductivity of

semiconductors can often be greatly increased by adding very small amounts of other elements. For example, the conducting ability of silicon or germanium can be immensely enhanced by adding traces of arsenic or boron to these elements. This property of semiconductors has many practical applications.

1.2 Coulomb's law

The *electrostatic force* of attraction or repulsion between two charges (strictly speaking point charges) was first measured quantitatively by Charles Augustine de Coulomb in 1875. The apparatus used by him is shown in Fig. 1.2.

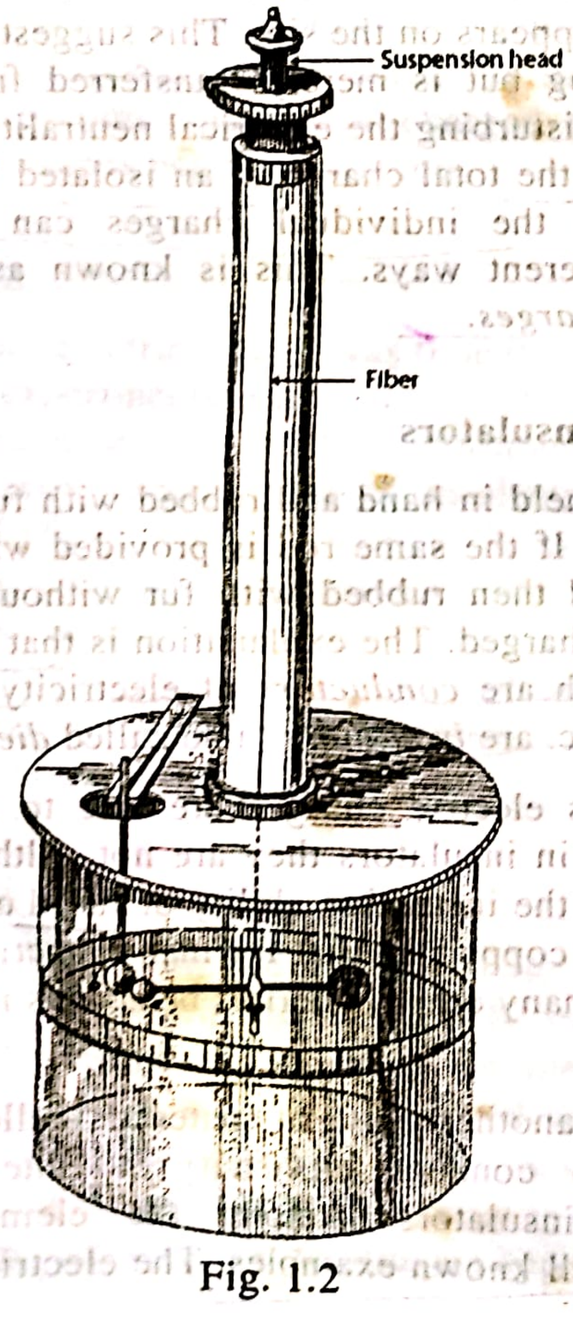


Fig. 1.2

The charges are confined to small spheres a and b. Of the two spheres b is fixed while a is suspended by a fibre. As b is fixed, the electric force on a will tend to twist the suspension fibre. Coulomb cancelled out this twisting effect by turning the suspension head through the angle θ needed to keep the two charges at a particular separation. The angle θ is then a relative measure of the electric force acting on a. The device of Fig. 1.2 is called a *torsion balance.*

Coulomb at first kept the magnitude of the charges on the spheres a and b constant and varied the distance of separation between the charges. His first experimental results can be represented by

$$F \propto \frac{1}{r^2}$$

where F is the magnitude of the force that acts on each of the two charges separated by a distance r. As required by Newton's third law of motion, these forces must act along the line joining the charges but point in opposite directions. It may be noted that the magnitude of the force on each charge is the same, even though the charges may be different.

Coulomb then studied how the electrical force varied with the relative size of the charges on the spheres of his torsion balance by keeping the distance fixed. If a charged conducting sphere is touched to an exactly similar but uncharged conducting sphere, the original charge will divide equally between the spheres. Coulomb varied the relative size of the charges by this technique and his experimental result can be represented by

$$F \propto q_1 q_2$$

where q_1 and q_2 are the charges on the two spheres.

Thus the electrostatic force exerted on one charge by another charge depends directly on the product of the magnitudes of the two charges and inversely on the square of their separation. That is

$$F \propto \frac{q_1 q_2}{r^2} \qquad (1.1)$$

Turning the above proportionality into an equation by introducing a constant of proportionality k, the force between two charges is given by

$$F = k\frac{q_1 q_2}{r^2} \tag{1.2}$$

Eqn. (1.2). is called *Coulomb's law* and generally holds for *point charges*. Charged objects whose sizes are much smaller than the distance between them are usually referred to as point charges.

The SI unit of charge is the *coulomb* (abbreviation C), which is defined as *the amount of charge* that *flows in one second when there is a steady current of one ampere*. That is

$$dq = idt \tag{1.3}$$

where dq (in coulombs) is the charge transferred by a current i (in amperes) during the interval dt (in seconds). For example, a current of 1A delivers a charge of 1×10^{-6} C in a time of 10^{-6} s.

In the SI system, the constant k is expressed in the following form:

$$k = \frac{1}{4\pi\epsilon_0}$$

where the constant ϵ_0, called the *permittivity constant* of free space (vacuum), must have that value which makes the right hand-side of eqn. (1.2) equal to the left-hand side. This value turns out to be

$$\epsilon_0 = 8.85418 \times 10^{-12} \text{ coul}^2 / \text{nt - m}^2.$$

The constant k has the corresponding value

$$k = \frac{1}{4\pi\epsilon_0} = 8.99 \times 10^9 \text{ N.m}^2/\text{coul}^2.$$

With this choice of k, Coulomb's law can be written as

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2} \tag{1.4}$$

When k has the above value, expressing q in coulombs and r in metres gives the force in newtons.

Coulomb's law has survived every experimental test; no exceptions to it have ever been found. The significance of Coulomb's law goes far beyond the description of the forces acting between charged balls or rods. This law, when incorporated into the structure of quantum physics, correctly describes (a) the electric forces that bind the electrons of an atom to its nucleus, (b) the forces

that bind atoms together to form molecules, and (c) the forces that bind atoms or molecules together to form solids or liquids. Most of the forces of our daily experience that are not gravitational in nature are electrical.

1.3 Gravitational force and electrical force

In eqn. (1.2), F is the magnitude of the force acting on either particle owing to the charge on the other, and q_1 and q_2 are the *magnitudes* (or *absolute values*) of the charges of the two particles. The gravitational force between the two particles is given by

$$F = G\frac{m_1 m_2}{r^2}$$

where m_1 and m_2 are the masses of the two particles, G being gravitational constant. By analogy with G, the constant k may be called the *electrostatic constant.* Both laws are inverse square laws and both involve a property of the interacting particles – the mass in one case and the charge in the other.

However there are dissimilarities between the laws too. The gravitational forces are always attractive but electrostatic forces may be either attractive or repulsive, depending on the signs of the two charges. This difference arises from the fact that, although there is only one kind of mass, there are two kinds of charges.

***Example 1.1** Determine the force between two free electrons spaced 1 Å (10^{-10}m) apart (a typical atomic dimension).*

Soln.

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}$$

$q_1 = q_2 = -1.6 \times 10^{-19}$C

$r = 1\,\text{Å} = 10^{-10}$m

$\frac{1}{4\pi\epsilon_0} = 9.0 \times 10^9$ N.m²/coul²

$$= \frac{(9.0\times10^9)(-1.6\times10^{-19})^2}{(1.0\times10^{-10})^2}$$

$$= 2.3 \times 10^{-8}\ \text{N}$$

$$= 23\ \text{nN}.$$

The force is repulsive.

***Example 1.2** The uranium nucleus contains a charge 92 times that of a proton. If a proton is shot at the nucleus, how large a repulsive force does the proton experience due to the nucleus when it is 1×10^{-11}m from the nucleus centre? The nuclei of atoms are of the order of the 10^{-14}m in diameter, so the nucleus can be considered a point charge.*

Soln.

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2} \qquad q_1 = 92.e = (92 \times 1.6 \times 10^{-19}C)$$

$$q_2 = 1.6 \times 10^{-19}C$$

$$= \frac{(9\times10^9)(92\times1.6\times10^{-19})(1.6\times10^{-19})}{(1\times10^{-11})^2}$$

$$= 2.1 \times 10^{-4}N.$$

***Example 1.3** Two point charges q_1 and q_2 are 3m apart, and their combined charge is 20μC. What are the magnitude of the two charges if (a) one repels the other with a force of 0.075N, and (b) attracts the other with a force of 0.525N?*

Soln.

(a) Since the force is repulsive, we have from the relation

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2},$$

$$0.075 = (9 \times 10^9)\frac{q_1 q_2}{r^2}$$

or, $q_1 q_2 = \dfrac{0.075\times9}{9\times10^9} = 75 \times 10^{-12} C^2 = 75\mu C^2$.

Now $q_1 + q_2 = 20\mu C$, or, $q_2 = (20 - q_1)$

$\therefore \quad q_1(20 - q_1) = 75;$

or, $q^2 - 20q_1 + 75 = 0;$ or, $q_1^2 - 5q_1 - 15q_1 + 75 = 0$

or, $q_1(q_1 - 5) - 15(q_1 - 5) = 0$, or, $(q_1 - 5)(q_1 - 15) = 0$

or, $q_1 = 5$ or, $15\,\mu C$

Hence $q_2 = 15$ or, $5\mu C$.

(b) Force is attractive; hence

$$-0.525 = (9 \times 10^9)\frac{q_1 q_2}{9}$$

$$\text{or, } q_1 q_2 = \frac{(-0.525)(9)}{9\times10^9} = -525\mu C^2.$$

Again substituting $q_2 = (20 - q_1)$

$$q_1(20 - q_1) = 525$$

or, $q_1^2 - 20q_1 - 525 = 0$

or, $q_1^2 + 15q_1 - 35q_1 - 525 = 0$

or, $q_1(q_1 + 15) - 35(q_1 + 15) = 0$

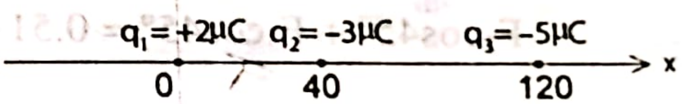
or, $(q_1 + 15)(q_1 - 35) = 0$

or, $q_1 = -15$ or, $+35\mu C$ and $q_2 = +35$ or, $-15\mu C$.

Example 1.4 Three point charges are placed at the following points on the x-axis: +2μC at x = 0, –3μC at x = 40cm. –5μC at x = 120cm. Find the force on the –3μC charge.

Soln.

The arrangement is shown in the adjoining figure. The force on q_2 is the vector sum of two contributions, the attractive force due to q_1 and the repulsive force due to q_3.



The direction of both the forces is towards the left *i.e.*, towards q_1. The sum of these two forces, taken algebraically, since they are along the same line, is

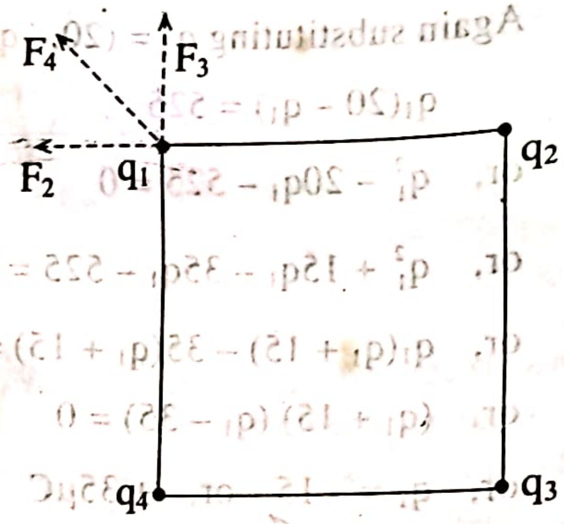
$$F = F_1 + F_2 = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_1^2} + \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r_2^2}$$

$$= 9 \times 10^9 \left[\frac{(2\times10^{-6})(3\times10^{-6})}{(0.40)^2} + \frac{(5\times10^{-6})(3\times10^{-6})}{(0.80)^2}\right]$$

$= 0.55$N towards left.

Example 1.5 *Four equal point charges, +3mC, are placed at the four corners of a square that is 40 cm on a side. Find the force on any one of the charges.*

Soln.



The situation is depicted in the adjoining figure. Let us consider the forces acting on q_1. Let F_2, F_3 and F_4 be the forces exerted on q_1 by q_2, q_3 and q_4 respectively.

By symmetry

$$F_2 = F_3 = (9 \times 10^9)\frac{(3\times10^{-6})(3\times10^{-6})}{(0.40)^2} = 0.51 \text{ N}.$$

Since the directions of F_2 and F_3 are along the edges as shown in the figure, their vector sum will lie along the diagonal from q_4 to q_1 and have the magnitude.

$$F_2\cos45° + F_3\cos45° = 0.51\frac{1}{\sqrt{2}} + 0.51\frac{1}{\sqrt{2}}$$

$$= (2)(0.51)\frac{1}{\sqrt{2}} = 0.72\text{N}.$$

The remaining force F_4 is also along this diagonal, and

$$F_4 = (9 \times 10^9)\frac{(3\times10^{-6})(3\times10^{-6})}{\left(\sqrt{0.40^2+0.40^2}\right)^2} = 0.25 \text{ N}$$

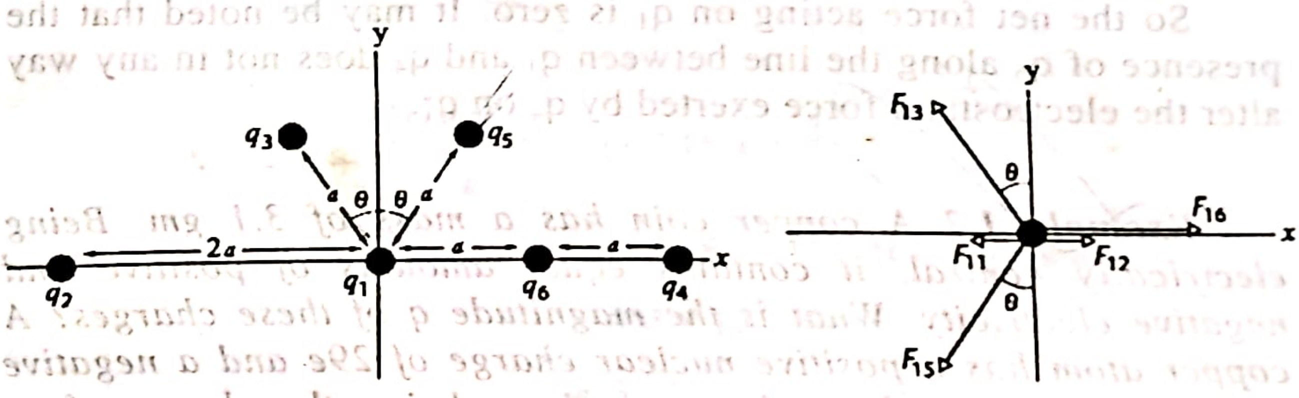
The resultant of all the three force acting along the diagonal away from the square.

$$= 0.72 \text{ N} + 0.25 \text{ N} = 0.97 \text{ N}.$$

Example 1.6 The figure below shows an arrangement of six fixed charged particles, where a = 2.0 cm and θ = 30°. All six particles have the same magnitude of charge q = 3.0 x 10^{-6}C; their electrical signs are as indicated. What is the net electrostatic force F_1 acting on q_1 due to other charges?

Soln.

From the figure it is obvious that F_1, the net electrostatic force acting on q_1 is the vector sum of F_{12}, F_{13} F_{14}, F_{15} and F_{16}, the exerted on q_1 due to other charges. Because q_2 and q_4 are equal in magnitude and are both a distance r = 2a from q_1, we have from eqn. (1.4)



(a) (b)

$$F_{12} = F_{14} = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{(2a)^2} \quad \text{(i)}$$

Similarly, since q_3, q_5 and q_6 are equal in magnitude and are each a distance r = a from q_1, we have

$$F_{13} = F_{15} = F_{16} = \frac{1}{4\pi\epsilon_0}\frac{q_1q_3}{a^2} \quad \text{(ii)}$$

Fig. (b) is a free body diagram for q_1. It and eqn. (i) show that F_{12} and F_{14} are equal in magnitude but opposite in direction, thus they cancel each other. Inspection of the figure (b) and eqn. (ii) reveals that the y-components of F_{13} and F_{15} also cancel and that their x-components are identical in magnitude and both point in the direction of decreasing x. Fig. (b) also shows that F_{16} points in the direction of increasing x. Thus F_1 must be parallel to the x-axis; its magnitude is the difference between F_{16} and twice the x-component of F_{13}.

$$\therefore \quad F_1 = F_{16} - 2F_{13} \sin\theta$$

$$= \frac{1}{4\pi\epsilon_0} \frac{q_1 q_6}{a^2} - 2.\frac{1}{4\pi\epsilon_0} \frac{q_1 q_3}{a^2} \sin\theta$$

Now $q_3 = q_6$ and $\theta = 30°$,

Substituting these values we find

$$F_1 = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_6}{a^2} - 2.\frac{1}{4\pi\epsilon_0} \frac{q_1 q_6}{a^2} \sin 30° = 0$$

So the net force acting on q_1 is zero. It may be noted that the presence of q_6 along the line between q_1 and q_4 does not in any way alter the electrostatic force exerted by q_4 on q_1.

Example 1.7 *A copper coin has a mass of 3.1 gm. Being electrically neutral, it contains equal amounts of positive and negative electricity. What is the magnitude q of these charges? A copper atom has a positive nuclear charge of 29e and a negative electronic charge of equal magnitude, e being the charge of an electron.*

Soln.

The number N of copper atoms in the coin

$$N = \frac{m}{M}.N_o$$

where m = mass of the coin

M = atomic weight of copper

N_o = Avogadro's number

$$= \frac{3.1}{64} \times 6.0 \times 10^{23}$$

$= 2.9 \times 10^{22}$ atoms.

Charge (either positive or negative) of a copper atom

$$= 29 \times 1.6 \times 10^{-19}C = 4.64 \times 10^{-18}C.$$

Therefore, the charge

$$q = (4.64 \times 10^{-18})\,(2.9 \times 10^{22}) = 1.345 \times 10^{5}C.$$

This is an enormous amount of charge. It can be verified that it will take 40 hrs. for a charge of this amount to pass through a 100 watt, 110 volt light bulb.

Example 1.8 *What should be the distance of separation between the total positive and negative charges of the copper coin of example 1.7 such that their force of attraction is 4.5N?*

Soln.

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2} \qquad q_1 = q_2 = 1.345 \times 10^{5}C.$$

$$F = 4.5\text{ N}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{q^2}{r^2}$$

$$\therefore \quad r = q.\sqrt{\frac{1/4\pi\epsilon_0}{F}} = (1.345 \times 10^{5})\sqrt{\frac{9\times10^{9}}{4.5}}$$

$$= (1.345 \times 10^{5})\,(0.4472 \times 10^{5})$$

$$= 6.0 \times 10^{9} \text{ metres}$$

$$= 3.75 \times 10^{6} \text{ miles.}$$

Example 1.9 *The distance r between the electron and the proton in the hydrogen atom is about 5.3×10^{-11} metre. What are the magnitudes of (a) the electrical force and (b) the gravitational force between these two particles?*

Soln.

$$F_e = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2} \qquad q_1 = q_2 = 1.6 \times 10^{-19}C$$

$$r = 5.3 \times 10^{-11}m.$$

$$= (9 \times 10^9)\,\frac{(1.6\times10^{-19})^2}{(5.3\times10^{-11})^2}$$

$$= 8.1 \times 10^{-8}\ \text{N.}$$

The gravitational force is given by

$$F_g = G.\frac{m_1 m_2}{r^2}$$

$$= \frac{(6.7\times10^{-11})(9.1\times10^{-31})(1.7\times10^{-27})}{(5.3\times10^{-11})}$$

$$= 3.7 \times 10^{-47}\ \text{N.}$$

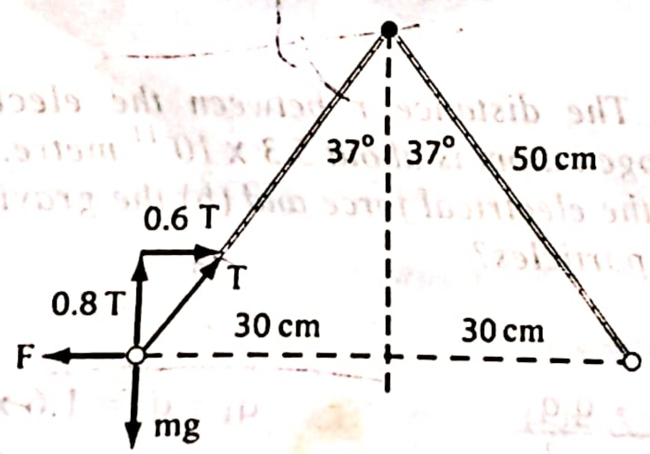
m_1 = electron mass = 9.1×10^{-31}Kg
m_2 = proton mass = 1.7×10^{-27}Kg
G = gravitational constant = 6.7×10^{-11}N-m²/kg²
r = 5.3×10^{-11}m.

Thus the electrical force is approximately 2.2×10^{39} times stronger than the gravitational force.

***Example 1.10** The two balls shown in the figure below have identical masses of 0.20gm each. When suspended from 50cm long strings, they make an angle of 37° to the vertical. If the charges on each are the same, how large is each charge?*

Soln.

Since the system is at rest, let us apply the conditions for equilibrium to the ball on the left. Three forces act on this ball: its weight mg; the tension T in the string; and F, the repulsive force due to the charge on the other ball.



Applying the usual conditions for equilibrium:

$$\Sigma F_x = 0\,; \qquad i.e.,\ F - 0.6T = 0 \qquad \text{(i)}$$

$$\Sigma F_y = 0\,; \qquad i.e.,\ 0.8T - mg = 0$$

$$\text{or,}\quad 0.8T - (0.2 \times 10^{-3}\text{kg})\,(9.8\text{m/s}^2) = 0$$

$$\text{or,}\quad T = 2.45 \times 10^{-3}\text{N} \qquad \text{(ii)}$$

Substituting the value of T as given by (ii) in (i) we get

$$F - (0.6)\,(2.45 \times 10^{-3}\text{N}) = 0$$

$$\text{or,}\quad F = 1.47 \times 10^{-3}\text{N} \qquad \text{(iii)}$$

Substituting this value of F in Coulomb' law

$$1.47 \times 10^{-3} = (9 \times 10^{9})\,\frac{q^2}{(0.60)^2}$$

$$\text{or,}\quad q^2 = \frac{(1.47\times10^{-3})(0.60)^2}{9\times10^{9}} = 0.0588 \times 10^{-12}$$

$$= 5.88 \times 10^{-14}$$

$$\text{or,}\quad q = 2.4248 \times 10^{-7}\text{C}$$

$$= 0.24248 \times 10^{-6}\text{C}$$

$$= 0.24248\mu\text{C}.$$

1.3 Electric field and electric field strength

Suppose that a stationary positive charge q_o, placed at any point in a certain region of space, experiences a force **F**. We then say there is an *electric field* in the region, whose strength **E** at the point is given by

$$\mathbf{E} = \frac{\mathbf{F}}{q_0} \qquad (1.5)$$

The *electric field strength* **E**, also known as *electric intensity*, at a point is, therefore, defined as the *force per unit positive charge* (assumed positive for convenience). Here **E** is a vector because **F** is

one, q_0 being scalar. The direction of **E** is the direction of **F**, that is, it is the direction in which a stationary positive charge placed at the point would tend to move. The condition that q_0 be stationary is necessary to distinguish electric field from magnetic field as moving charges give rise to magnetic fields. It is evident from eqn. (1.5) that the unit of electric field strength is *Newton per coulomb.*

If **E** itself is produced by a set of charges *fixed in space*, then its value as given by eqn. (1.5) is independent of the size of q_0 and depends on the positions and magnitudes of other charges. On the other hand, **E** may in practice be produced by charges on bodies such as conductors. In that case, when the charge q_0 is introduced, it may cause a redistribution of the source charges and the value of **E** will then depend on the size of q_0. While we might be interested in the field **E** for large values of q_0, the most useful concept is that of the electric field which would exist at a point due to the original undisturbed charges. This field is defined by making the magnitude of q_0, usually referred to as the *test charge*, as small as possible so that it does not disturb the field produced by the source charges. The electric field is, therefore, defined as

$$\mathbf{E} = \lim_{q_0 \to 0} \frac{\mathbf{F}}{q_0} \tag{1.6}$$

q_0 cannot of course be smaller than the electronic charge. Unless otherwise specified, any value quoted for E will be that for an undisturbed source.

Example 1.11 *What is the magnitude of the electric field strength E such that an electron, placed in the field, would experience an electrical force equal to its weight?*

Soln.

$$\mathbf{E} = \frac{\mathbf{F}}{q_0} = \frac{mg}{q_0}$$

$$= \frac{(9.1\times10^{-31})(9.8)}{1.6\times10^{-19}}$$

$$= 5.6 \times 10^{-11} \text{ N/C}$$

$m = 9.1 \times 10^{-31}$ kg

$g = 9.8 \text{ m/s}^2$

$q_0 = 1.6 \times 10^{-19}$ C

Lines of force

The concept of the electric field as a vector was not appreciated by Michael Faraday. In order to visualize the electric field in space, he introduced the concept of lines of force. The lines of force are supposed to originate from the positive charges and terminate in the negative charges. These lines may be straight or curved depending on the system of charge creating them. The relationship between the (imaginary) lines of force and the electric field strength vector is given below.

1. The tangent to the line of force at any point gives the *direction* of **E** at that point.

2. The density of the lines of force (the number of lines of force crossing unit area perpendicular to the field) is proportional to the magnitude of **E**. Where the lines are close together **E** is large and where they are far apart **E** is small.

No two lines of force can cross each other; because in that case there will be two directions of the electric field strength at the point of cross-over which is impossible. Lines of force are used for visualizing electric field patterns – they are not employed quantitatively.

Calculation of electric field strength

Let a test charge q_0 be placed at a distance r from a point charge. The magnitude of the force acting on q_0 as given by Coulomb's law is

$$F = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2}$$

The strength of the electric field at the site of the test charge as given by eqn. (1.5) is

$$E = \frac{F}{q_0} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \qquad (1.7)$$

The direction of **E** is on a radial line from q, pointing outward if q_0 is positive and inward if q is negative.

If there are more than one charge in the field then (i) the electric field strength at the given point should be calculated for

each charge *as if it were the only charge present* and (ii) these separately calculated fields should be added vectorially to find the resultant field strength **E** at the point. In equation form,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \ldots\ldots = \sum \mathbf{E}_n \quad n = 1, 2, 3, \ldots \tag{1.8}$$

The sum is a vector sum, taken over all the charges. If the charge distribution is not discrete but continuous, then the charge must be divided into infinitesimal elements of charge dq. The field dE due to each element of charge at the point in question is then calculated. The magnitude of dE is given by

$$dE = \frac{1}{4\pi\epsilon_0} \frac{dq}{r^2} \tag{1.9}$$

where r is the distance of the given point from the charge element dq. The resultant field at the point is then obtained by adding vectorially the field contributions due to all charge elements, but in most cases the addition needs to be carried out by integration.

or, $\mathbf{E} = \int d\mathbf{E}$

In carrying out the integration, the only suitable method is to resolve the field contributions from the charge elements into components, add the components by integration giving say, $\mathbf{E_x}$, $\mathbf{E_y}$ and $\mathbf{E_z}$ from which **E** can be obtained.

Field due to a uniformly charged wire (line of charge):

Let us consider a section of an infinite line of charge having a charge density (*i.e.*, the charge per unit length) λ coulomb per metre-oriented as shown in Fig. 1.3. We would like to calculate the electric field **E** at a point P a distance y from the line.

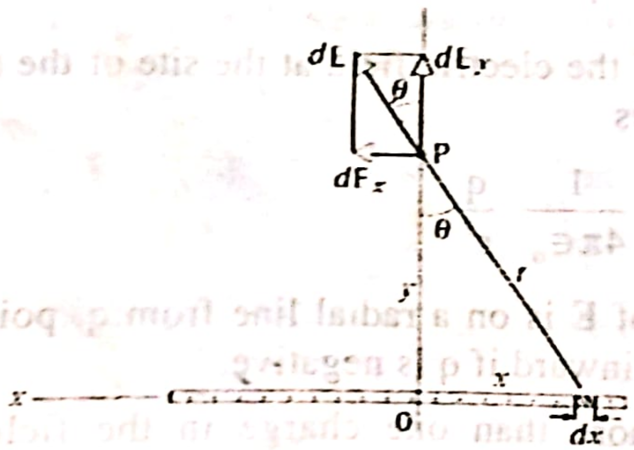


Fig. 1.3

The magnitude of the field contribution dE at the point P due to a charge element dq (= λ.dx) is given, using eqn. 1.8, by

$$dE = \frac{1}{4\pi\epsilon_0} \frac{dq}{r^2}$$

where r is the distance of the point P from the charge element dq.

Now $r^2 = x^2 + y^2$

$$\therefore \quad dE = \frac{1}{4\pi\epsilon_0} \frac{dq}{x^2 + y^2} = \frac{1}{4\pi\epsilon_0} \frac{\lambda.dx}{x^2 + y^2} \qquad (1.10)$$

The vector dE can be resolved, as shown in the figure, into two components

$$dE_x = -\,dE \sin\theta \quad \text{and} \quad dE_y = dE \cos\theta$$

The minus sign in front of dE_x indicates that dE_x points in the negative x direction.

The x and y components of the resultant vector E at point P is given by

$$E_x = \int dE_x = -\int_{x=-\infty}^{x=+\infty} \sin\theta dE$$

$$\text{and} \quad E_y = \int dE_y = \int_{x=-\infty}^{x=+\infty} \sin\theta dE$$

If the wire is infinitely long and be symmetrically placed on both right and left of the perpendicular PO, then for any element of charge dq = λ.dx to the right of O, there is an equal element of charge in a symmetrical position to the left of O. Consequently, the field contribution in the x direction made by these two symmetric elements of charge cancel each other. E_x must, therefore, be zero. Thus E points entirely in the y-direction. Because the contributions to E_y from the right and left halves are equal, E_y, and hence E, may be written as

$$E = E_y = 2\int_{x=0}^{x=+\infty} \sin\theta.dE \qquad (1.11)$$

It may be noted that the lower limit of the integration has been changed and a factor of 2 introduced.

From Fig. 1.3,

$$x = y \tan\theta \qquad \text{so that} \quad dx = y \sec^2\theta . d\theta$$

$$r = y/\cos\theta = y\sec\theta \quad \text{so that} \quad r^2 = x^2 + y^2 = y^2\sec^2\theta$$

Now $dE_y = dE \cos\theta$

$$= \left(\frac{1}{4\pi\epsilon_0} . \frac{\lambda . dx}{x^2 + y^2}\right) \cos\theta$$

$$= \frac{1}{4\pi\epsilon_0} \frac{\lambda . y \sec^2\theta d\theta}{y^2 \sec^2\theta} \cos\theta$$

$$= \frac{1}{4\pi\epsilon_0 y} \cos\theta . d\theta$$

Thus,

$$E = 2 \int_{x=0}^{x=+\infty} \cos\theta . dE = 2 \int_{x=0}^{x=+\infty} \frac{\lambda}{4\pi\epsilon_0 y} \cos\theta . d\theta$$

$$= \frac{2\lambda}{4\pi\epsilon_0 y} \int_{\theta=0}^{\theta=\pi/2} \cos\theta d\theta$$

as $x = 0$ corresponds to $\theta = 0$ and $x = +\infty$ corresponds to $\theta = \pi/2$.

$$\therefore \quad E = \frac{\lambda}{2\pi\epsilon_0 y} \left|\sin\theta\right|_0^{\pi/2}$$

$$= \frac{\lambda}{2\pi\epsilon_0 y} \qquad (1.12)$$

Any actual line must have a finite length – not an infinite length. However, for points close enough to finite lines and far from their ends, eqn. (1.11) yields results that are so close to the correct values that the difference can be ignored in many practical situations.

(vii) Field at a point on the axis of a charged circular ring.

Let us consider a circular turn of wire of radius *a* carrying a charge q. We would like to calculate E at a point P on the axis of the ring a distance x from its centre.

The arrangement is shown in Fig. 1.4. Let us consider a differential element of the ring of length ds at the top of the ring. Then the charge contained in this element is given by

$$dq = \frac{ds}{2\pi a} . q$$

where $2\pi a$ is the circumference of the ring. This element produces a field dE at the point P.

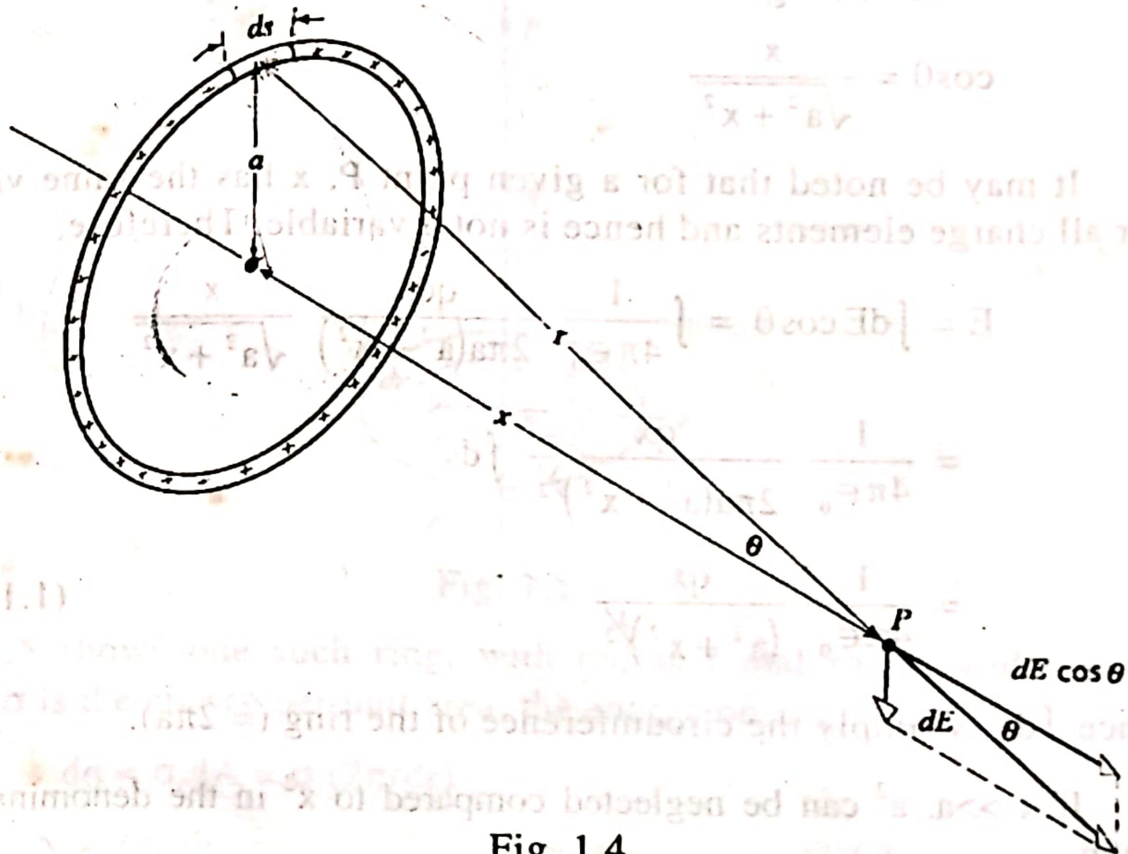


Fig. 1.4

The resultant field **E** at P is found by integrating the effects of all the elements that make up the ring.

or, $\mathbf{E} = \int d\mathbf{E}$

The component of dE perpendicular to the axis is cancelled out by an equal opposite component established by the charge element

on the opposite side of the ring. Thus only the component of dE parallel to the axis of the ring contributes to the resultant field.

The general vector integral

$$E = \int dE$$

becomes a scalar integral $E = \int dE \cos\theta$

Now dE is given by

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dq}{r^2} = \frac{1}{4\pi\epsilon_0}\left(\frac{qds}{2\pi a}\right)\cdot\frac{1}{a^2+x^2}$$

where $r^2 = a^2 + x^2$

Also from the figure we have

$$\cos\theta = \frac{x}{\sqrt{a^2+x^2}}$$

It may be noted that for a given point P, x has the same value for all charge elements and hence is not a variable. Therefore,

$$E = \int dE\cos\theta = \int \frac{1}{4\pi\epsilon_0}\frac{qds}{2\pi a(a^2+x^2)}\frac{x}{\sqrt{a^2+x^2}}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{qx}{2\pi a(a^2+x^2)^{3/2}}\int ds$$

$$= \frac{1}{4\pi\epsilon_0}\frac{qx}{(a^2+x^2)^{3/2}} \qquad (1.13)$$

since $\int ds$ is simply the circumference of the ring ($= 2\pi a$).

If $x \gg a$, a^2 can be neglected compared to x^2 in the denominator. Then

$$E = \frac{1}{4\pi\epsilon_0}\cdot\frac{q}{x^2} \qquad (1.14)$$

and is the same as that produced by a point charge.

Thus greater the distance of the point from the charged ring, the more nearly does the field produced by the ring approach that produced by a point charge.

(iii) Electric field due to a charged disk.

A circular plastic disk of radius R that has a positive surface charge of uniform density σ on its upper surface is shown in Fig. 1.5. We would like to calculate the electric field at point P, a distance x from the disk along its central axis. To do so the disk is first divided into concentric flat rings and then the electric field at point P is calculated by adding up *i.e.*, by integrating the contributions of all the rings.

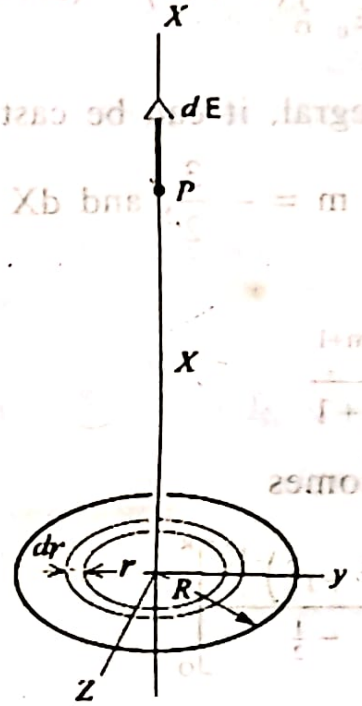


Fig. 1.5

Fig. 1.5 shows one such ring, with radius r and radial width dr. Since σ is the charge per unit area, the charge on the ring is

$$dq = \sigma.dA = \sigma\,(2\pi r dr)$$

where $dA = 2\pi r.dr$ ... (i)

The electric field dE at point P due to this flat ring, as already given by eqn. 1.13, is

$$dE = \frac{1}{4\pi\epsilon_0}\frac{x.(\sigma.2\pi r dr)}{\left(r^2+x^2\right)^{3/2}} \qquad \text{(ii)}$$

where a in eqn. 1.13 has been replaced by r.

Eqn. (ii) may be written as

$$dE = \frac{\sigma x}{4\epsilon_0}\frac{2rdr}{(r^2+x^2)^{3/2}} \qquad \text{(iii)}$$

E can now be obtained by integrating eqn. (ii) over the surface of the disk, *i.e.*, by integrating with respect to the variable r from r = 0 to r = R, x remaining constant during the process.

$$E = \int dE = \frac{\sigma x}{4\epsilon_0}\int_0^R (x^2+r^2)^{-3/2}\,(2r)\,dr \qquad \text{(iv)}$$

To solve this integral, it can be cast in the form $\int X^m dX$ by setting $X = (r^2 + x^2)$, $m = -\frac{3}{2}$, and $dX = (2r)\,dr$. For the recast integral, we have

$$\int X^m dX = \frac{X^{m+1}}{m+1}$$

Eqn. (iii) then becomes

$$E = \frac{\sigma . x}{4\epsilon_0}\left[\frac{(x^2+r^2)^{-1/2}}{-\frac{1}{2}}\right]_0^R$$

Taking the limits and rearranging, we find

$$E = \frac{\sigma}{2\epsilon_0}\left(1 - \frac{x}{\sqrt{x^2+R^2}}\right) \qquad (1.15)$$

as the magnitude of the electric field produced by a flat, circular disk on its central axis. In carrying out the integration it has been assumed that $x \geq 0$.

If we allow R to become infinity ($R \to \infty$) while keeping x finite, the second term in the parenthesis in eqn. (1.15) approaches zero and the equation reduces to

$$E = \frac{\sigma}{2\epsilon_0} \quad \dots\dots\dots \text{(infinite sheet)} \qquad (1.16)$$

This is the electric field produced by an infinite sheet of uniform charge, located on one side of a nonconductor such as plastic.

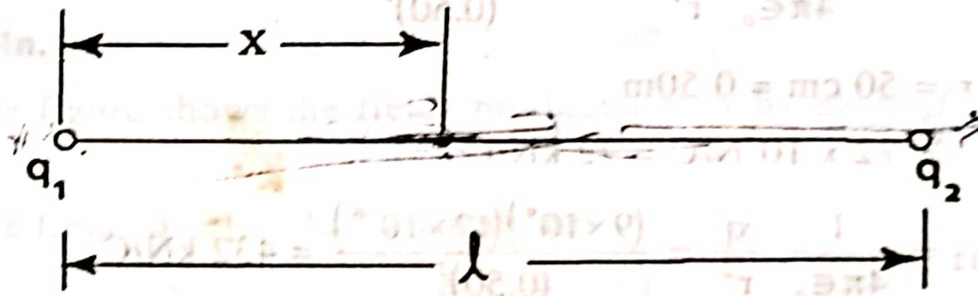
If we let $x \to 0$ in eqn. (1.15) while keeping R finite, we again obtain (1.16). This shows that at points very close to the disk, the electric field set up by the disk is the same as if the disk were infinite in extent. At points very far from the disk, the electric field produced by the disk is the same as if it were a point charge.

***Example 1.12** Figure shows a charge $q_1(=+1.0 \times 10^{-6}$ coul) 10 cm from a charge q_2 (= + 2.0 x 10^{-6} coul). At what point on the line joining the two charges is the electric field strength zero?*

Soln.

The point must lie between the charges because only here do the forces exerted by q_1 and q_2 on the test charge oppose each other.

Let the point P be at a distance x from q_1. If E_1 and E_2 are the electric field strengths at P due to q_1 and q_2 respectively, we must have



$$E_1 = E_2$$

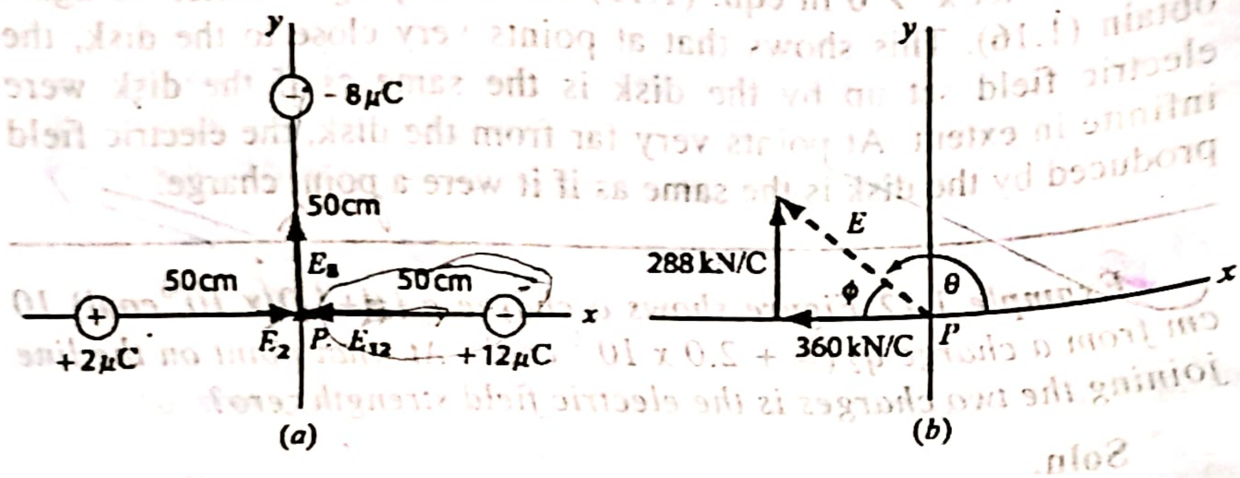
or, $\dfrac{1}{4\pi\epsilon_0}\dfrac{q_1}{x^2} = \dfrac{1}{4\pi\epsilon_0}\dfrac{q_2}{(l-x)^2}$ where l = 10 cm.

or, $\dfrac{(l-x)^2}{x^2} = \dfrac{q_2}{q_1} = 2;$ or, $\dfrac{l-x}{x} = \sqrt{2}$

or, $\dfrac{10}{x} - 1 = \sqrt{2};$ or, $\dfrac{10}{x} = 1+\sqrt{2}$

or, $x = \dfrac{10}{1+\sqrt{2}} = 4.1$ cm.

Example 1.13 Find electric field at point P in figure below due to the charges shown.



Soln.

The field due to each charge are shown in the figure.

$$E_2 = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = \frac{(9\times10^9)(2\times10^{-6})^2}{(0.50)^2}$$

where $r = 50$ cm $= 0.50$m.

$$= 72 \times 10^3 \text{N/C} = 72 \text{ kN/C}$$

$$E_{12} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = \frac{(9\times10^9)(12\times10^{-6})}{(0.50)^2} = 432 \text{ kN/C}$$

$$E_8 = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = \frac{(9\times10^9)(8\times10^{-6})}{(0.50)^2} = 288 \text{ kN/C}$$

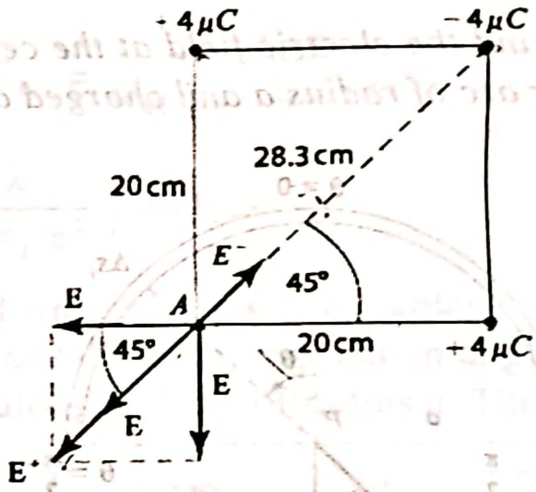
Of these, E_2 and E_{12} act along the x-axis in opposite directions. Their resultant is (432 – 72) = 360 kN/C which act along the negative x-axis, as shown in Fig. (b). E_8 acts along the positive y-axis, also shown in the figure. Their resultant is

$$E = (E_2 + E_{12}) + E_8$$

$$= 461 \text{ kN/C}$$

The resultant field acts in a direction making an angle of 39° with the –ve x-axis $[\cos\phi = \frac{288}{360}$; $\phi = 39°]$.

Example 1.14 *Three charges are placed at three corners of a square of side 20 cm as shown in the figure below. Find the magnitude and direction of the electric field strength at point A.*



Soln.

The figure shows the fields produced at A by each of the three charges.

The fields due to +4μC charges are

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = \frac{(9\times10^9)(4\times10^{-6})}{(0.20)^2}$$

$$= 0.9 \text{ MN/C}.$$

Their directions are shown in the figure. Their resultant is

$$E^+ = \sqrt{(9\times10^5)^2 + (9\times10^5)^2} = 1.27 \text{ MN/C}.$$

and its direction is as shown in the figure. The field due to the negative charge is

$$E^- = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} = \frac{(9\times10^9)(-4\times10^{-6})}{\left[\sqrt{(0.20)^2 + (0.20)^2}\right]^2} = -0.45 \text{ MN/C}$$

and its direction is as shown.

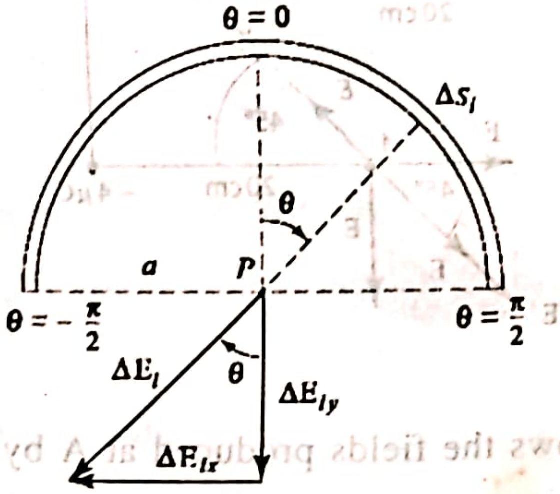
Both E^+ and E^- act along the same line but in opposite directions. Hence the net field at the point A is

$$E_t = E^+ + E^- = 0.82 MN/C$$

and its direction is away from the negative charge as shown in the figure.

***Example 1.15** Find the electric field at the centre of a uniformly charged semicircular arc of radius a and charged density of λ C/m.*

Soln.



Let the arc be split into small segments ΔSi. The segments are sufficiently small to be considered to be a point charge λΔSi. The electric field due to a segment at the centre of the arc *i.e.*, at point P, is given by

$$\Delta E_i = \frac{1}{4\pi \epsilon_0} \frac{\lambda \Delta Si}{a^2}$$

Each little segment of the arc will give a ΔE_i in a different direction. The resultant field at the point P is the vector sum of the fields due to all the segments.

To add these fields vectorially, the fields shoud be resolved into components. The resultant component in the x-direction, E_x, will be zero at the point P. This is because the ΔE_{ix} shown in the figure will be cancelled by the contribution from a symmetrically placed ΔS on the left half of the arc. As a result, we need only compute E_y at point

P in order to find the resultant field at the point P. As can be seen from the figure, the direction of the field is vertically downwards *i.e.*, along the – ve y-direction.

Now

$$\Delta E_{iy} = \frac{1}{4\pi\epsilon_0}\frac{\lambda\Delta Si}{a^2}\cos\theta$$

Integrating,

$$E_y = \int \Delta E_{iy} = \int \frac{1}{4\pi\epsilon_0}\frac{\lambda\Delta Si}{a^2}\cos\theta$$

$$= \frac{\lambda}{4\pi\epsilon_0 a^2}\int \cos\theta\, ds$$

The integrand involves two variables, θ and s. s can be expressed in terms of θ by recalling that an angle dθ subtends an arc of length ds = adθ along a circle of radius a. Therefore,

$$E_y = \frac{\lambda}{4\pi\epsilon_0 a^2}\int_{-\pi/2}^{+\pi/2} \cos\theta . a d\theta$$

$$= \frac{\lambda}{4\pi\epsilon_0 a}\int_{-\pi/2}^{+\pi/2} \cos\theta\, d\theta$$

$$= \frac{\lambda}{4\pi\epsilon_0 a}\left|\sin\theta\right|_{-\pi/2}^{+\pi/2}$$

$$= \frac{\lambda}{2\pi\epsilon_0 a}.$$

1.4 A point charge in an electric field

A particle of charge q in an electric field **E** experiences a force **F** given by (eqn. 1.5)

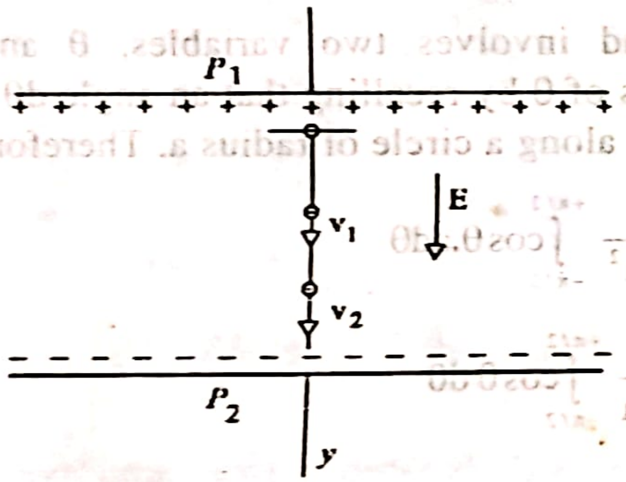
F = qE

The force will produce an acceleration

a = F / m

where m is the mass of the particle. Two examples of the acceleration of a charged particle in a uniform electric field will be considered. Such a uniform electric field can be produced by connecting the terminals of a battery to two parallel metal plates which are otherwise insulated from each other. If the spacing between the plates is small compared with the dimensions of the plates, the field between them will be fairly uniform except near the edges. In describing the motion of the charged particle in a field set up by external charges, the *self-field, i.e.*, the field due to the particle itself is ignored.

(i) A particle of mass m and charge q is placed at rest in a uniform electric field and released. The resulting motion resembles that of a body falling in the earth's gravitational field (Fig. 1.6). The (constant)



Ex. 1.6

acceleration is given by

$$a = \frac{F}{m} = \frac{qE}{m}$$

The equations for uniformly accelerated motion then apply. With the uniform velocity $v_0 = 0$, we have

$$v = v_0 + at = at = \frac{qEt}{m} \quad \text{(i)}$$

$$y = v_0 t + \frac{1}{2}at^2 = \frac{1}{2}at^2 = \frac{qEt^2}{2m} \quad \text{(ii)}$$

and $$v^2 = v_0^2 + 2ay = \frac{2qEy}{m} \qquad \text{(iii)}$$

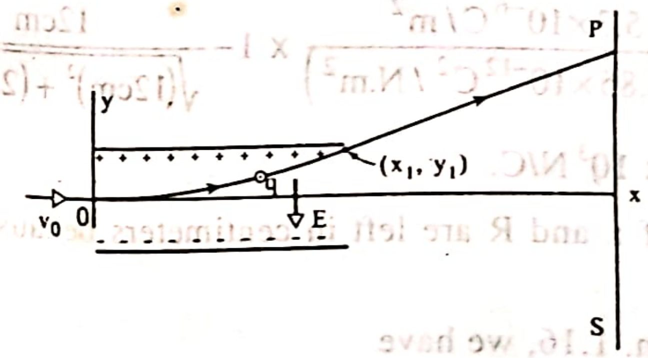
The kinetic energy attained after falling though a distance y is found from

$$K = \frac{1}{2} mv^2 = \frac{1}{2} m \left(\frac{2qEy}{m}\right)$$

$$= qEy \qquad \text{(iv)}$$

Since a constant force qE acts over a distance y, the result given by eqn. (iv) can be obtained directly from work-energy theorem.

(ii) An electron of mass m and charge e is projected with speed v_0 at right angles to a uniform electric field. The resulting motion is like that of a projectile fired horizontally in the earth's gravitational field (Fig. 1.7). The horizontal (x) and the vertical (y) components of motion are given by



$$x = v_0 t \qquad \text{(v)}$$

and $$y = \frac{1}{2} at^2 = \frac{eE}{2m} t^2 \qquad \text{(vi)}$$

(since the initial vertical component is zero)

Eliminating t, we get

$$y = \frac{eE}{2mv_0^2} x^2 \qquad \text{(vii)}$$

for the equation of the trajectory.

On emerging from the plates, the electron travels in a straight line tangent to the parabola at the exit point. If the electron is allowed to fall on a fluorescent screen S placed some distance beyond the plates, then together with other electrons following the same path, it will make itself as a small luminous spot. This is the principle of the *cathode-ray oscilloscope.*

***Example 1.16** A disk has a radius R of 2.5 cm and a surface charge density σ of $+5.8\mu C/m^2$ on its upper face. Calculate (i) the electric field at a point on the central axis at a distance x = 12 cm from the disk and (ii) the electric field at the surface of the disk.*

Soln.

(i) From eqn. 1.15, we have

$$E = \frac{\sigma}{2\epsilon_0}\left(1 - \frac{x}{\sqrt{x^2 + R^2}}\right)$$

$$= \frac{5.3\times10^{-6}C/m^2}{(2)(8.85\times10^{-12}C^2/N.m^2)} \times 1 - \frac{12cm}{\sqrt{(12cm)^2 + (2.5cm)^2}}$$

$= 6.3 \times 10^3$ N/C.

The values of x and R are left in centimeters because that unit cancels.

(ii) From eqn. 1.16, we have

$$E = \frac{\sigma}{2\epsilon_0} = \frac{5.3\times10^{-6}C/m^2}{(2)(8.85\times10^{-12}C^2/N.m^2)}$$

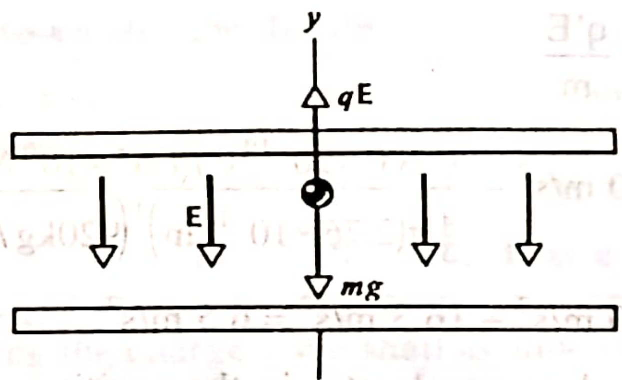
$= 3.0 \times 10^5$ N/C.

***Example 1.17** A charged drop of oil of radius R = 2.76μm and density p = 920 kg/m^3 is maintained in equilibrium under the combined influence of its weight and a downward uniform electric field of magnitude E = 1.65 x 10^6 N/C as shown in the figure. (i) Calculate the magnitude and sign of the charge on the drop. Express the result in terms of the elementary charge e. (ii) The drop is exposed to a radioactive source that emits electrons. Two electrons*

strike the drop and are captured by it, changing its charge by two units. If the electric field remains unchanged, calculate the resulting acceleration of the drop.

Soln.

The weight of the drop acts downwards. So, to keep the drop in equilibrium, the electric force of magnitude qE must act upwards. Since the direction of the field is downward, the charge q on the drop must be negative for the electric force to point in a direction opposite the field. For the condition of equilibrium, the resultant force on the drop



$$\sum F = mg + qE = 0$$

Taking y components, we get

$$-mg + q(-E) = 0$$

$$\text{or,}\quad q = -\frac{mg}{E} = \frac{\frac{4}{3}\pi R^3 \rho . g}{E} \qquad (\because m = \frac{4}{3}\pi R^3 \times \rho)$$

$$= \frac{-\frac{4}{3}\pi (2.76\times10^{-6}\,m)^3 (920 kg/m^3)(9.8 m/s^2)}{1.65\times10^6\,N/C}$$

$$= 4.8 \times 10^{-19}\ C.$$

Writing q in terms of the electronic charge – e as q = n(– e) where n is the number of electronic changes on the drop, we have

$$n = \frac{q}{-e} = \frac{-4.8\times10^{-19}C}{-1.6\times10^{-19}C} = 3.$$

(ii) If two additional electrons are added to the drop, then its charge will become

$$q' = (n + 2)(-e) = 5 \times (-1.6 \times 10^{-19}\,C) = -8.0 \times 10^{-19}C.$$

Applying Newton's second law of motion, the resultant force on the drop is equal to

$$\Sigma F = mg + q'E = ma$$

Taking the y-components, we get

$$-mg + q'(-E) = ma$$

$$\text{or,} \quad a = -g - \frac{q'E}{m}$$

$$= -9.80\ m/s^2 - \frac{(-8.0\times10^{-19}C)(1.65\times10^{6}N/C)}{\frac{4}{3}\pi(2.76\times10^{-16}m)^3(920kg/m^3)}$$

$$= -9.80\ m/s^2 + 16.3\ m/s^2 = 6.5\ m/s^2.$$

Therefore, the drop accelerates in the positive y-direction, *i.e.*, upwards.

***Example 1.18** The electric field between the plates of a cathode-ray oscilloscope is 1.2 x 10⁴N/C. What deflection will an electron experience if it enters at right angles to the field with a kinetic energy of 2000 eV(= 3.2 x 10⁻¹⁶ joules), a typical value. The deflecting assembly is 1.5 cm long.*

Soln.

Recalling that $k_0 = \frac{1}{2}mv_0^2$, eqn. (iii) ($y = \frac{eE}{2mv_0^2}x^2$)

can be written as

$$y = \frac{eEx^2}{4k_0}$$

If x_1 is the horizontal position of the far edge of the plate, y_1 will be the corresponding deflection (Fig. 1.11),

$$\text{or,} \quad y_1 = \frac{eEx_1^2}{4k_0}$$

$$= \frac{(1.6\times10^{-19}\,C)(1.2\times10^{4}\,N/C)(1.5\times10^{-2}\,m)^2}{(4)(3.2\times10^{-16}\,\text{joule})}$$

$$= 3.4 \times 10^{-4}\text{m} = 0.34 \text{ mm.}$$

the deflection measured, not at the deflecting plates but at the fluorescent screen, is much larger.

1.5 Field due to an electric dipole

An electric dipole consists of two equal but opposite charges separated by a small distance. Fig. 1.8 shows an electric dipole. The charges are +q and −q and they are separated by a distance 2a.

We would like to calculate the field **E** at a point P due to the dipole. P is located at a distance x along the perpendicular bisector of the line joining the charges. We shall assume that x >> a.

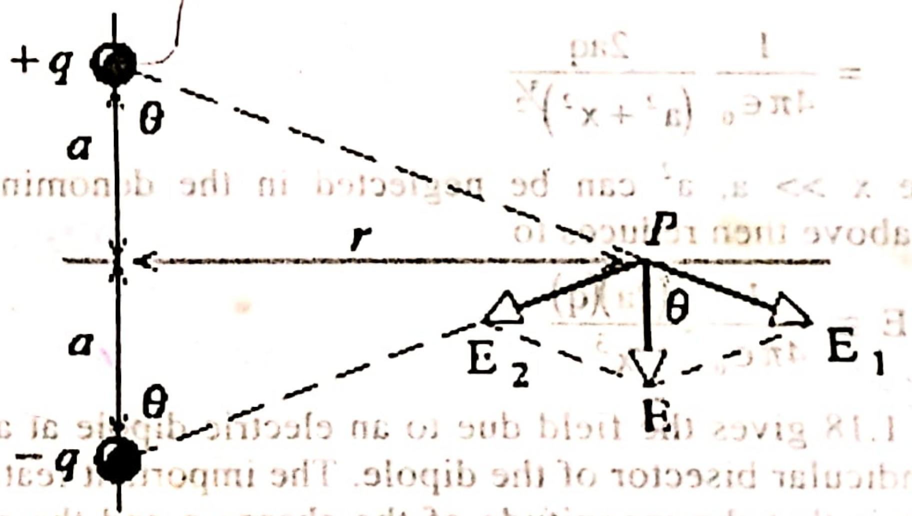


Fig. 1.8

Let the distance of the point P be y from both the charges. Let E_1 be the electric field at the point P due to the charge +q and E_2 be that due to −q. The total field at P due to the dipole is obtained by vector addition of

$$\mathbf{E} = \mathbf{E_1} + \mathbf{E_2}$$

Since the charges have the same magnitude and the distance of the point P from the charges is also same, the magnitudes of the fields E_1 and E_2 are equal. Or,

$$E_1 = E_2 = \frac{1}{4\pi\epsilon_0}\frac{q}{y^2} = \frac{1}{4\pi\epsilon_0}\frac{q}{a^2+x^2}$$

The directions of E_1 and E_2 are indicated by arrows in Fig. 1.8. The horizontal components of E_1 and E_2 cancel each other. Hence the vector sum of E_1 and E_2 points vertically downwards and has the magnitude

$$E = 2E_1 \cos\theta$$

From the figure

$$\cos\theta = \frac{a}{y} = \frac{a}{\sqrt{a^2+x^2}}$$

Substituting the values of E_1 and $\cos\theta$ in eqn. 1.17 we obtain

$$E = \frac{2}{4\pi\epsilon_0}\frac{q}{(a^2+x^2)}\cdot\frac{a}{\sqrt{a^2+x^2}}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{2aq}{(a^2+x^2)^{3/2}}$$

since x >> a, a^2 can be neglected in the denominator. The equation above then reduces to

$$E \cong \frac{1}{4\pi\epsilon_0}\cdot\frac{(2a)(q)}{x^3} \qquad (1.18)$$

Eqn. 1.18 gives the field due to an electric dipole at a point on the perpendicular bisector of the dipole. The important feature of the expression is that the magnitude of the charge q and the separation 2a between the charges enter eqn. 1.18 only as a product. This means that if we measure **E** at various distances from the electric dipole (assuming x >> a), we can never deduce '*q*' and '*a*' separately but only the product 2aq; if q were doubled and 'a' simultaneously cut in half, the electric field at large distance from the dipole would not change.

The product 2aq is called the *electric dipole moment* and is denoted by *p*. It is a vector having the direction along the axis of the dipole from the negative to the positive charge.

Another important feature of eqn. 1.18 hat the field E varies with distance x as $1/x^3$. This implies that when two equal and opposite charges are placed close to each other, their separate fields at distant points almost, *but not quite*, cancel each other; whereas for a point charge, E(x) drops off more slowly, namely as $1/x^2$ (eqn. 1.7).

1.6 A dipole in an external electric field

An electric dipole is placed in a uniform external electric field **E**; its dipole moment **p** making an angle θ with this field [Fig. 1.9(a)]. The two forces (**F** and – **F**) acting on the charges are equal and opposite where

$$F = qE$$

The net force on the dipole is clearly zero. But since the forces do not act along the same line, there is a net torque on the dipole about an axis passing through the centre O of the dipole given by

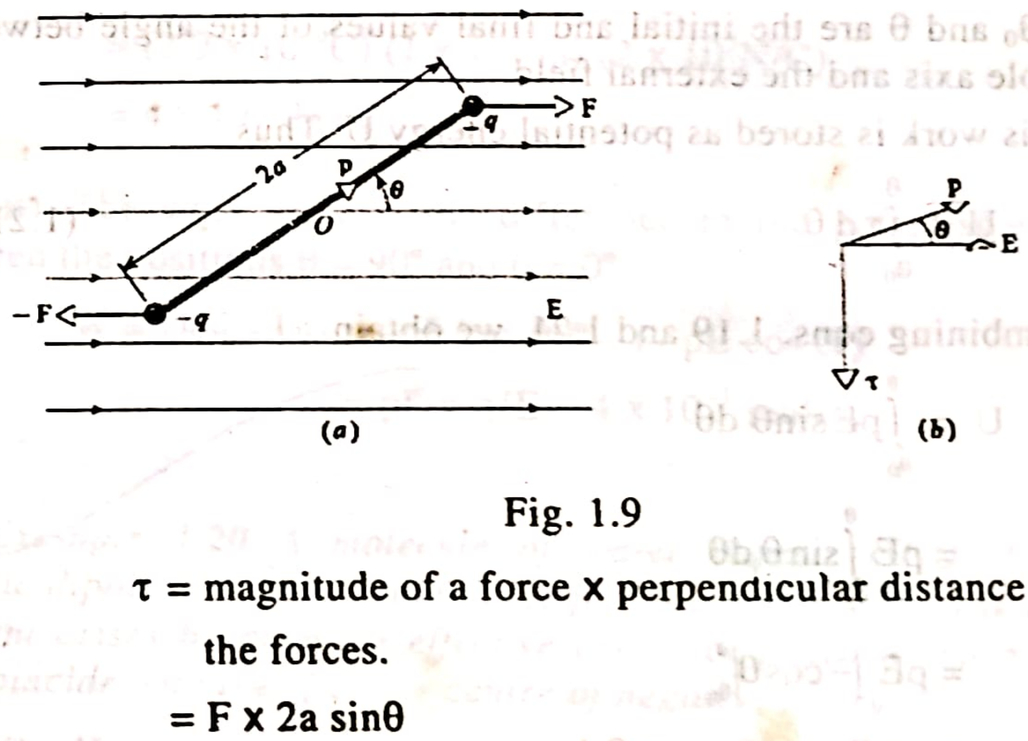


Fig. 1.9

τ = magnitude of a force x perpendicular distance between the forces.

$= F \times 2a \sin\theta$

$= qE \times 2a \sin\theta$

$= 2aq\ E \sin\theta$

$= p\ E \sin\theta \qquad (p = 2aq) \qquad (1.19)$

eqn. 1.19 can be written in vector form as

$$\tau = \mathbf{p} \times \mathbf{E} \tag{1.20}$$

the appropriate vectors being shown in Fig. 1.9(b).

Thus, when an electric dipole is placed in an external electric field, it experiences a torque tending to align it with this field. Hence, work (positive or negative) must be done by an external agent to change the orientation of an electric dipole in an external electric field. This work is stored as potential energy U in the system consisting of the dipole and the arrangement used to set up the external field. The work done to change the orientation of the dipole through a small angle dθ is given by

$$dW = \tau . d\theta$$

where τ is the torque exerted by the agent and does the work.

Then the work done to turn the dipole from an initial orientation θ_0 to a final orientation θ is given by

$$W = \int_{\theta_0}^{\theta} dW = \int_{\theta_0}^{\theta} \tau . d\theta$$

where θ_0 and θ are the initial and final values of the angle between the dipole axis and the external field.

This work is stored as potential energy U. Thus

$$U = \int_{\theta_0}^{\theta} \tau . d\theta \tag{1.21}$$

Combining eqns. 1.19 and 1.21, we obtain

$$U = \int_{\theta_0}^{\theta} pE \sin\theta \, d\theta$$

$$= pE \int_{\theta_0}^{\theta} \sin\theta \, d\theta$$

$$= pE \left| -\cos\theta \right|_{\theta_0}^{\theta}$$

$$= -\, pE \,(\cos\theta - \cos\theta_0)$$

Since we are interested only in *changes* in potential energy, the reference orientation θ_0 can be chosen to have any convenient value, say 90º. This gives,

$U = -pE\cos\theta$

This, again, can be put in the vector from

$U = -\mathbf{p}\cdot\mathbf{E}$

Example 1.19 An electric dipole consists of two opposite charges of magnitude 2.0×10^{-6}C separated by a distance $l = 1.0$cm. It is placed in an external electric field of 2.0×10^{5} N/C.

(i) What maximum torque does the field exert on the dipole?

(ii) How much work must an external agent do to turn the dipole from its initial alignment given by $\theta = 0°$ to final alignment $\theta = 90°$?

Soln.

(i) The maximum torque is found by putting $\theta = 90°$ in eqn. 1.19

$\tau = pE\sin\theta$ $q = 2.0 \times 10^{-6}$C

$l = 1.0 \text{ cm} = 1 \times 10^{-2}$m

$= qlE\sin 90°$ $E = 2.0 \times 10^{5}$ N/C

$= (2.0 \times 10^{-6}\text{C})(1 \times 10^{-2}\text{m})(2 \times 10^{5}\text{N/C})$

$= 4 \times 10^{-3}$ N.m.

(ii) The work done is the difference in the potential energy between the positions $\theta = 90°$ and $\theta = 0°$.

$W = U_{90°} - U_{0°} = -pE\cos 90° - (-pE\cos 0°)$

$= pE = qlE = 4 \times 10^{-3}$ joules.

Example 1.20 A molecule of water vapour (H_2O) has an electric dipole moment of magnitude $p = 6.2 \times 10^{-30}$C.m. The dipole moment arises because the effective centre of positive charge does not coincide with the effective centre of negative charge.

(i) How far apart are the effective centres of positive and negative charge in a molecule of H_2O?

(ii) What is the maximum torque on a molecule of H_2O in an electric field of magnitude 1.5×10^4 N/C?

(iii) Suppose the dipole moment of a molecule of H_2O is initially pointing in a direction opposite to the field. How much work is done by the electric field in rotating the molecule into alignment with the field?

Soln.

(i) There are 10 electrons and, correspondingly, 10 positive charges in this molecule. We can write, for the magnitude of the dipole moment,

$$p = qd = (10e)\,d$$

where d is the separation between the centres of the positive and negative charges. e is the elementary charge. Thus

$$d = \frac{p}{10e} = \frac{6.2\times10^{-30}\,C.m}{(10)\left(1.6\times10^{-19}\,C\right)}$$

$$= 3.9 \times 10^{-12} m = 3.9 \text{ pm}$$

(ii) As in eqn. 1.19, the torque is maximum when $\theta = 90°$. Substituting this value in that equation yields,

$$\tau = pE \sin\theta$$

$$= (6.2 \times 10^{-30}\,C.m)\,(1.5 \times 10^{4}\,N/C)\,(\sin 90°)$$

$$= 9.3 \times 10^{-26}\,N.m.$$

(iii) The work done in rotating the dipole from $\theta_0 = 180°$ to $\theta = 0°$ is given by eqn.

$$W = \int_{\theta_0}^{\theta} pE \sin\theta . d\theta$$

$$= \left| -pE\cos\theta \right|_{180°}^{0°}$$

$$= -pE \cos 180° - (-pE \cos 0°)$$

$$= pE + pE$$

$$= 2pE = (2)\,(6.2 \times 10^{-30} C.m)\,(1.5 \times 10^{4}\,N/C)$$

$$= 1.9 \times 10^{-25}\,J.$$

EXERCISES

1. State and explain Coulomb's law of electrostatics.

2. Discus the similarities and dissimilarities between gravitational and electrical forces. Show that the electrical force between the electron and proton in the hydrogen atom is approximately 2.2 x 10^{39} times stronger than the gravitational forces between them.

3. What are electric field and electric field strength? Derive an expression for the electric field due to a point charge.

4. Show that the electric field due to an infinite line of charge at a point is given by $E = \lambda/(2\pi\epsilon_0 y)$ where λ is the linear charge density and y is the perpendicular distance of the point from the line.

5. Calculate the electric field at a point on the axis of a charged circular ring. Hence show that the greater the distance of the point from the charged ring, the more nearly does the field produced by the ring approach that produced by a point charge.

6. Show by direct integration that the magnitude of the electric field due to an infinite plane sheet of charge is $\sigma/2\epsilon_0$, where σ is the surface charge density.

7. What is an electric dipole? Calculate the electrical field due to an electric dipole at a point on the perpendicular bisector of the dipole. What is electric dipole moment?

8. An electric dipole is placed in a uniform external electric field **E**. Show that (i) the torque exerted on the dipole by the field is given by $\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}$ and (ii) the work that must be done to change the orientation of the dipole in the field is given by $W = -\mathbf{p} \cdot \mathbf{E}$ where **p** is the electric dipole moment.

9. What repulsive Coulomb force exists between two protons in a nucleus of iron? Assume a separation of 4.0 x 10^{-15}m. **[14 N]**

10. The uranium nucleus contains a charge 92 times that of the proton. If a proton is shot at the nucleus, how large a repulsive force does the proton experience due to the nucleus when it is 1 x 10^{-11}m from the nucleus centre? The nuclei of atoms are of the order 10^{-14}m in diameter, so the nucleus can be considered a point charge. **[2.1 x 10^{-4}N]**

11. Four equal-magnitude (4 μC) charges are placed at the four corners of a square that is 20 cm on each side. Find the electric-field intensity at the centre of the square, if the charges are all positive **[E = 0]**.

12. Two charges of $+1\mu C$ and $-1\mu C$ are placed at the corners of the base of an equilateral triangle. The length of a side of the triangle is 0.7m. Find the electric-field intensity at the apex of the triangle. **[18.4 KN/C to the right].**

13. An electric dipole consists of two opposite charges of magnitude $q = 1.0 \times 10^{-6}$C separated by d = 2.0cm. The dipole is placed in an external field of 1.0×10^{5} N/C. (i) What maximum torque does the field exert on the dipole? (ii) How much work must an external agent do to turn the dipole end for end, starting from a position of alignment ($\theta = 0°$)? **[2.0×10^{-3}N.m; 4.0×10^{-3} joule]**

14. An electron is shot at 10^{6} m/s between two parallel charged plates. If E between the plates is 1 KN/c, where will the electron strike the upper plate? Assume vacuum conditions. **[The electron travels 0.75cm in the horizontal direction before hitting the plate].**

15. An oil carries six electronic charges, has a mass of 1.6×10^{-12}g, and falls with a constant velocity in air. What magnitude of vertical electric field is required to make the drop move upward with the same speed as it was formerly moving downward? [32.7 KN/C]

16. Two point charges q_1 and q_2 of unknown magnitude and sign are distance r apart. The electric field strength is zero at a distance r/3 from q_1 on the line joining them. What can you conclude about the charges?

17. Calculate the electric field at a distance of one metre on the axis of a circular turn of wire of radius 2 cm carrying charges of 20 coulombs per metre of its length.

CHAPTER II

GAUSS' LAW AND ITS APPLICATIONS

Flux of the Electric Field

Flux is the property of any vector field and is usually denoted by the symbol φ. It refers to a hypothetical surface in the field which may be closed or open. For an *electric field* the flux φ_E is measured by the *number of lines of force that cut through such a surface.* For closed surfaces the flux φ_E is considered *positive* if the lines of force point *outward* everywhere and negative if they point *inward.*

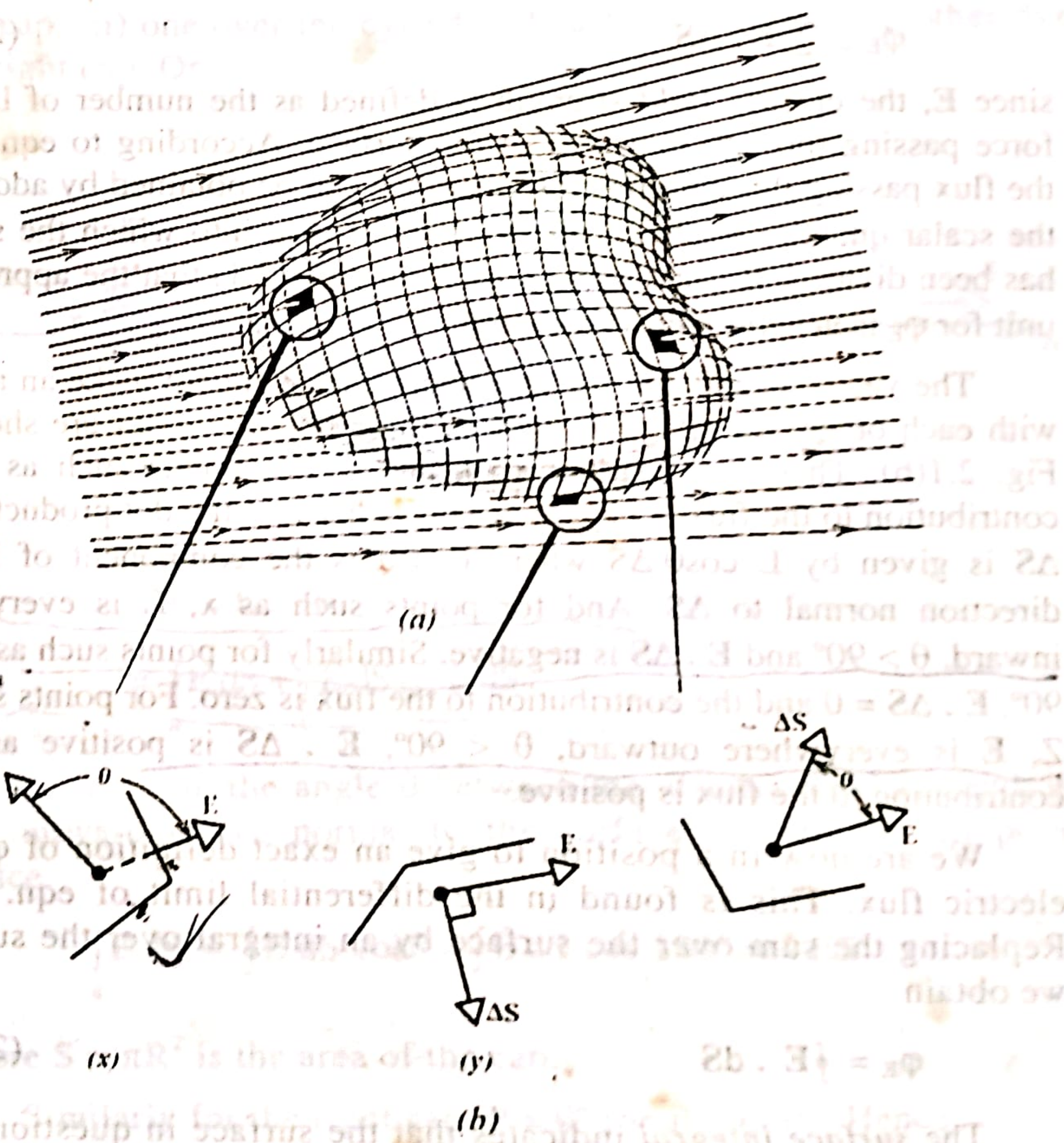


Fig. 2.1

For a precise definition of φ_E, consider a closed surface of arbitrary shape immersed in an electric field as shown in Fig. 2.1. The surface may be imagined to be divided into a large number of elementary squares ΔS, each of which is small enough so as to be considered to be a plane. Such an element of area can be represented as vector $\Delta \mathbf{S}$, whose magnitude is the area ΔS, the direction of ΔS being taken as the *outward drawn normal* to the surface.

An electric field vector **E** can also be constructed at every square. Since the squares have been taken to be arbitrarily small, **E** may be taken as constants for all points on the square.

From a semi-quantitative definition of flux we can write

$$\varphi_E \simeq \sum \mathbf{E} \cdot \Delta \mathbf{S} \tag{2.1}$$

since **E**, the electric field strength is defined as the number of lines of force passing through unit area of the surface. According to eqn. (2.1), the flux passing through the closed surface can be obtained by adding up the scalar quantity $\mathbf{E}.\Delta\mathbf{S}$ for all elements of area into which the surface has been divided. It can also be seen from eqn. (2.1) that the appropriate unit for φ_E is *newton-metre²/coulomb*.

The vector **E** and ΔS that characterize each square make an angle θ with each other. An enlarged view of three possible cases are shown in Fig. 2.1(b). These are marked x, y, and z. For points such as x, the contribution to the flux is negative. This is because the dot product of $\mathbf{E} \cdot \Delta S$ is given by $E \cos\theta \, \Delta S$ where $E\cos\theta$ is the component of E in a direction normal to ΔS. And for points such as x, **E** is everywhere inward, $\theta > 90°$ and $\mathbf{E} \cdot \Delta S$ is negative. Similarly for points such as y, $\theta = 90°$, $\mathbf{E} \cdot \Delta S = 0$ and the contribution to the flux is zero. For points such as Z, **E** is everywhere outward, $\theta < 90°$, $\mathbf{E} \cdot \Delta S$ is positive and the contribution to the flux is positive.

We are now in a position to give an exact definition of φ_E, the electric flux. This is found in the differential limit of eqn. (2.1). Replacing the sum over the surface by an integral over the surface, we obtain

$$\varphi_E = \oint \mathbf{E} \cdot d\mathbf{S} \tag{2.2}$$

The *surface integral* indicates that the surface in question must first be divided into infinitesimal elements of area dS. The scalar

quantity E . dS must be evaluated for each element which should then be algebraically added to obtain the sum for the entire surface. The circle on the integral sign indicates that the surface over which the integration is to be carried out is a closed surface.

Example 2.1 A hypothetical cylinder of radius R is immersed in a uniform electric field E, the cylinder axis being parallel to the field. Determine φ_E for this closed surface.

Soln.

Fig. 2.2 shows the cylinder immersed in the electric field. The flux φ_E can be written as the sum of three integrals, (i) one over the left cap, (ii) one over the cylindrical surface, and (iii) the other over the right cap. Or

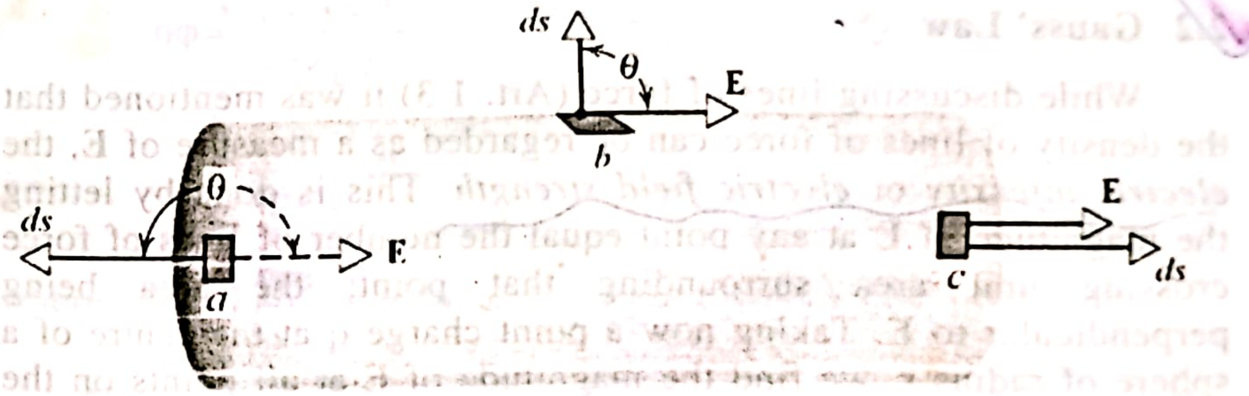


Fig. 2.2

$$\varphi_E = \oint .\mathbf{E} . d\mathbf{S}$$

$$= \int_a \mathbf{E}.d\mathbf{S} + \int_b \mathbf{E}.d\mathbf{S} + \int_c \mathbf{E}.d\mathbf{S}$$

For the left cap, the angle θ between the electric field strength and the outward drawn normal to the surface is 180° for all points. Hence

$$\int_a \mathbf{E}.d\mathbf{S} = \int E\, dS \cos\theta = \int E\, dS \cos 180^\circ = -ES$$

where $S = \pi R^2$ is the area of the cap.

Similarly for the right cap, $\theta = 0°$ for all points. Hence

$$\int_c \mathbf{E} \cdot d\mathbf{S} = \int_c E\, dS \cos 0^\circ = +ES$$

For the cylinder wall, $\theta = 90^\circ$ for all points on the cylindrical surface. Hence

$$\int_b \mathbf{E} \cdot d\mathbf{S} = \int_b E\, dS \cos 90^\circ = 0$$

Thus

$$\varphi_E = \int_a \mathbf{E} \cdot d\mathbf{S} + \int_b \mathbf{E} \cdot d\mathbf{S} + \int_c \mathbf{E} \cdot d\mathbf{S}$$

$$= -ES + 0 + ES$$

$$= 0.$$

2.2 Gauss' Law

While discussing lines of force (Art. 1.3) it was mentioned that the density of lines of force can be regarded as a measure of E, the *electric intensity* or *electric field strength*. This is done by letting the magnitude of E at any point equal the number of lines of force crossing unit area surrounding that point, the area being perpendicular to E. Taking now a point charge q at the centre of a sphere of radius r, we find the magnitude of E at all points on the surface to be given by $q/(4\pi\epsilon_0 r^2)$. Since the area of the spherical surface is $4\pi r^2$, this means that q/ϵ_0 lines of force leave a charge q.

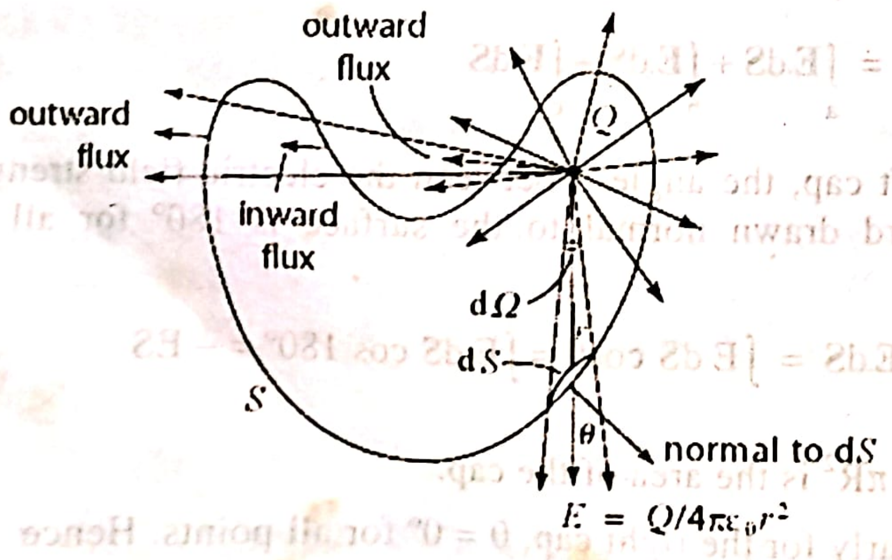


Fig. 2.3

The idea outlined above will now be put into more formal term by using, instead of the density of lines, the flux of the vector E. Let a point charge q be completely enclosed by an imaginary surface S (Fig. 2.3). An element dS of this surface is at a distance r from q. The electric field strength E at dS is $q/(4\pi\epsilon_0 r^2)$ in the direction of r. Hence, $d\varphi$, the flux of E across dS is given by

$$d\varphi = \mathbf{E} \cdot d\mathbf{S}$$
$$= E \cos\theta \, dS$$

where E cosθ is the component of **E** normal to the surface dS, θ being the angle between **E** and the outward drawn normal to the surface dS. (Alternately, dS cosθ is the projection of the area dS normal to the direction of the electric field strength **E**). Or,

$$d\varphi = \frac{q \, dS \cos\theta}{4\pi\epsilon_0 r^2}$$

$$= \frac{q}{4\pi\epsilon_0} d\Omega$$

where $d\Omega = \frac{dS\cos\theta}{r^2}$ is the solid angle subtended by dS at q.

Now the total solid angle subtended by S at q is 4π. Hence, the total flux of **E** over **S** is given by

$$\varphi_E = \oint \mathbf{E} \cdot d\mathbf{S}$$

$$= \oint \frac{q \, d\Omega}{4\pi\epsilon_0}$$

$$= \frac{q}{4\pi\epsilon_0} \oint d\Omega$$

$$= \frac{q}{4\pi\epsilon_0} \cdot 4\pi$$

$$= \frac{q}{\epsilon_0} \tag{2.3}$$

or, $$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q \tag{2.4}$$

The sign of the charge is to be incorporated in q (the inward flux being considered negative while the outward flux positive).

Because electric fields superpose, a number of charges inside S give their own fluxes each obeying eqn. (2.3). These fluxes add, giving a general theorem that *the outward flux of* **E** *over any closed surface is equal to the algebraic sum of the enclosed charges divided by* ϵ_0; or

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\epsilon_0} \sum q$$

$$\text{or, } \epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = \sum q \qquad (2.5)$$

Eqn. (2.5) is known as *Gauss' law*. Sometimes it is also referred to as *Gauss' theorem* (though not to be confused with a theorem of the same name in vector analysis).

(i) $\sum q$ in eqn. (2.5) is the *net* charge, taking its algebraic sign into account. *If a surface encloses equal and opposite charges, the flux* φ_E *is zero.*

(ii) *Charges outside the surface makes no contribution to the electric flux.* Consider a charge q lying outside the closed surface as shown in Fig. 2.4. The flux cone cut the area dS_1, at the point r_1 and

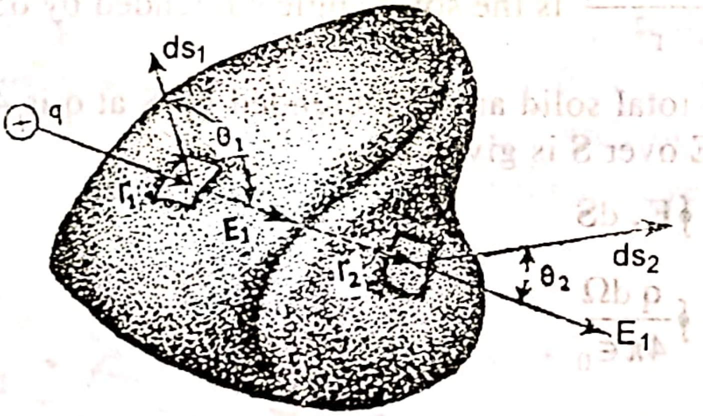


Fig. 2.4

the area dS_2 at r_2. The direction of the field due to the charge q is inward at r_1 on area dS_1 and outward at r_2 on area dS_2. But the solid angles subtended by dS_1 and dS_2 at q are the same. In other words,

$$d\Omega = \frac{dS_1 \cos\theta_1}{r_1^2} = \frac{dS_2 \cos\theta_2}{r_2^2}$$

$$\text{But } \mathbf{E}_1 . d\mathbf{S}_1 = -E_1 dS_1 \cos\theta_1 = -\frac{q}{4\pi\epsilon_0} \frac{dS\cos\theta_1}{r_1^2}$$

$$\text{and } \mathbf{E_2}.\mathbf{dS_2} = E_2 dS_2 \cos\theta_2 = \frac{q}{4\pi\epsilon_0} \frac{dS_2 \cos\theta_2}{r_2^2}$$

$$\text{Thus } \mathbf{E_1}.\mathbf{dS_1} + \mathbf{E_2}.\mathbf{dS_2} = \frac{q}{4\pi\epsilon_0}\left[\frac{dS_2 \cos\theta_2}{r_2^2} - \frac{dS_1 \cos\theta_1}{r_1^2}\right]$$

$$= \frac{q}{4\pi\epsilon_0}[d\Omega - d\Omega] = 0 \tag{2.6}$$

It is evident from Fig. 2.4 that the whole closed surface can be divided into pairs of such elementary areas and the positive and negative contribution on the pair due to a charge lying outside the closed surface will mutually cancel. Therefore, the sum total contribution due to it on the closed surface will be zero. Thus

$$\oint_S \mathbf{E}.\,\mathbf{dS} = 0 \tag{2.7}$$

if the charge q lies outside the closed surface.

(iii) *Also the total flux does not depend on the exact location of the charges inside the closed surface.*

If, instead of discrete charges, the free charge q is assumed to be a continuous distribution of charge throughout the entire volume enclosed by the closed surface, then we can write

$$\Sigma q = \int_V \rho.dv$$

where ρ is the volume density of the free charge in the element of volume dV, the integral being taken over the entire volume enclosed by the surface. Hence

$$\oint \mathbf{E}.\mathbf{dS} = \frac{1}{\epsilon_0}\int_V \rho.dV$$

Thus Gauss' law may also be stated as "the surface integral of the normal component of the electric field **E** over any closed surface or volume equals the total charge (or algebraic sum of charges) enclosed within the surface or equals $1/\epsilon_0$ times the net charge in free space within that volume."

Application of Gauss' law

(i) *Deduction of Coulomb's law*

Coulomb's law can be deduced from Gauss' law and symmetry considerations. To do so, let us apply Gauss' law to an isolated point charge q and consider that the charge is surrounded by a spherical surface as in Fig. 2.5. Although Gauss' law holds for any surface whatever, information can be extracted more readily for a spherical surface of radius r with the charge q at the centre. The advantage of this surface is that, from symmetry consideration the field **E** is normal to the surface and is constant in magnitude for all points on it. *A closed surface so imagined under symmetry consideration is called a Gaussian surface* – a term which we shall often use.

In Fig. 2.5 **E** and **dS** at any point on the Gaussian surface are directed radially outward. The angle between them is zero, and the quantity **E . dS** = E dS cosθ becomes EdS. Gauss' law then becomes

$$\epsilon_0 \oint \mathbf{E} . d\mathbf{S} = \epsilon_0 \oint EdS = q \qquad (2.8)$$

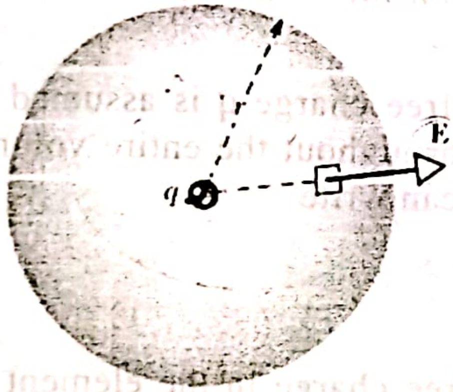


Fig. 2.5

Since E is constant for all points on the surface of the sphere, it can be taken outside the integral, and consequently eqn. 2.8 becomes

$$\epsilon_0 E \oint dS = q \qquad (2.9)$$

But $\oint dS$ is simply the area $4\pi r^2$ of the sphere. Eqn. (2.9) therefore reduces to

$$\epsilon_0 E 4\pi r^2 = q$$

$$\text{or, } E = \frac{q}{4\pi\epsilon_0 r^2} \qquad (2.10)$$

Eqn. (2.10) gives the magnitude of electric field strength **E** at any point a distance r from an isolated point charge q.

If now a second point charge q_0 is placed at the point where **E** has been calculated, the force acting on q_0 is

$F = Eq_0$

Combining this with eqn. (2.10) we obtain

$$F = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2}$$

which is nothing but the Coulomb's law of electrostatic force. Coulomb's law can, therefore, be deduced from Gauss' law and symmetry considerations.

(ii) *Electric field due to a charged sphere*

Let us consider a solid conducting sphere of radius R. if an amount of charge q is placed on this sphere, it will distribute itself uniformly over the surface of the sphere. *No charge can reside in the interior of the region*, because it will disturb the normal distribution of charges in this region of the conductor and hence will give rise to an unbalanced electric field causing an current to flow. Since there cannot be any such current in the case of electrostatic problems, *a conducting sphere must carry all the added charge on its surface.* The distribution of charge over the surface must also be uniform, otherwise there will exist a component of electric field tangential to the surface thus causing a current to flow on the surface. Again, this is not allowed under the assumed electrostatic condition. The distribution of change should be such that it would not create a tangential component of the field, so that the electric field **E** must be normal to the surface at all points on it. Since no tangential component of the field exists on the surface of the conductor, the surface will be an equipotential surface. The above discussion is true for any conductor of arbitrary shape. Consider a point P near but outside a uniform charged sphere of radius R with a charge q. Let us imagine a spherical Gaussian surface of radius r $(r > R)$ passing through the point P. The charge q is now surrounded by the Gaussian surface. By the spherical symmetry of the problem, if the electric field strength at P be E, then it must have the same magnitude at all points over the Gaussian surface and everywhere normal to it.

Further, since both **E** and d**S** at any point on the Gaussian surface are directed radially outwards, the angle between them is zero. So from Gauss' law, we have

$$\oint \mathbf{E}\cdot d\mathbf{S} = \oint E\cos 0^\circ dS = \oint E\, dS = \frac{q}{\epsilon_0}$$

Since E is uniform for all points on the surface, it can be brought outside the integral. Or,

$$E\oint dS = E.\,4\pi r^2 = \frac{q}{\epsilon_0}$$

$$\text{or, } E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \qquad (2.11)$$

Eqn. 2.11 is equivalent to the field at a point which is at a distance r from a point charge q. Thus *while finding the electric field outside a charged sphere, the charge can be considered to be concentrated at the centre of the sphere* and the formula for the point charge can be used.

Let us now consider a point inside the charged sphere and imagine a Gaussian surface of radius r (r < R) through the point. Applying Gauss' law we have

$$\oint E.\,dS = E\oint dS = E.4\pi r^2 = \frac{q'}{\epsilon_0}$$

where q' is the charge inside the Gaussian surface. But since all the charge resides on the surface of the charged sphere, there is no charge inside the sphere, *i.e.*, the charge inside the Gaussian surface is zero. Or, q' = 0. Hence the above relation reduces to

$$E\,.\,4\pi r^2 = 0$$

$$\text{or, } E = 0$$

Michael Faraday carried out an experiment designed to show that excess charge resides on the outside surface of a conductor. He built a large metal-covered box, which he mounted on insulating supports and charged it with a powerful electrostatic generator. In Faraday's words:

I went into the cube and lived in it, and using lighted candles, electrometers, and all other tests of electrical states, I could not find the least influence upon them ...... though all the time the outside of the cube was very powerfully charged, and large sparks and brushes were darting off from every part of its outer surface.

(ii) Field due to a spherically symmetric charge distribution (volume distribution of charge)

Fig. 2.6 shows a spherical region of radius R having a uniform volume distribution of charge. The charge density ρ, *i.e.*, the charge per unit volume (measured in coul/metre^3), at any point depends only on the distance of the point from the centre. This condition is

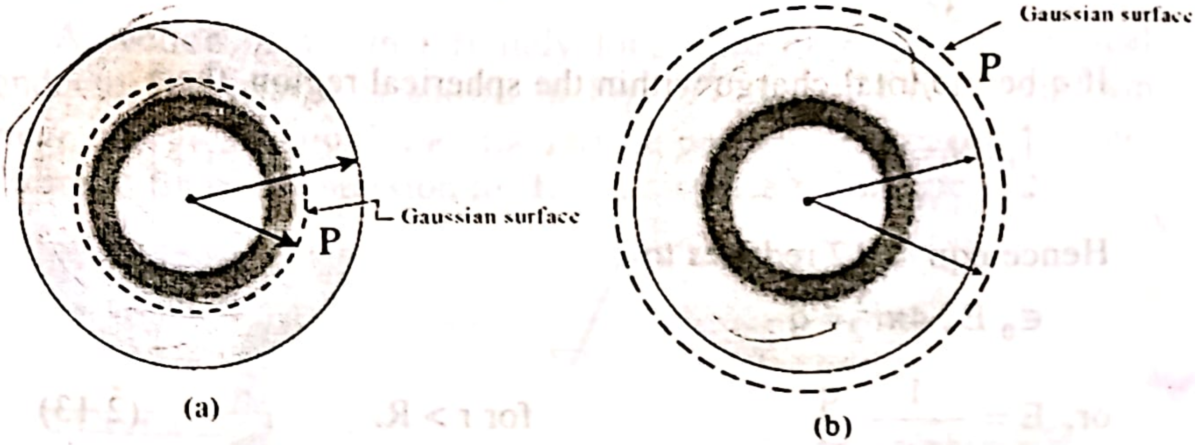


Fig. 2.6

referred to as spherical symmetry. Let us find an expression for E for points (b) outside and (a) inside the charge distribution. *The sphere of Fig. 2.6 cannot be a conductor because in that case all the charges given to the sphere i.e., the excess charges will reside on its surface.*

(a) *Field for points outside the sphere (r > R)*

Let us consider a point P outside the sphere at a distance r from the centre of the sphere. Consider a spherical Gaussian surface of radius r passing through the point. From symmetry and uniformity of the density, the field is radial and uniform over the Gaussian surface. Hence the angle between E and dS will be zero. Applying Gauss' law, we have

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = \epsilon_0 \oint E \cos 0° dS$$

$$= \epsilon_0 E 4\pi r^2 = \int_V \rho dV \qquad (2.12)$$

Since there is no charge outside the spherical region of charges, *i.e.*, ρ = 0 for r > R, the right side of eqn. 2.12 becomes

$$\int_V \rho dV = \rho . \frac{4}{3}\pi R^3$$

Hence eqn. 2.12 gives

$$E.4\pi r^2 = \frac{1}{\epsilon_0}\rho.\frac{4\pi}{3}R^3$$

$$\text{or, } E = \frac{\rho}{3\epsilon_0}\frac{R^3}{r^2} \quad \text{for } r > R.$$

If q be the total charge within the spherical region, then

$$\int_V \rho dV = q$$

Hence eqn. 2.12 reduces to

$$\epsilon_0 E.4\pi r^2 = q$$

$$\text{or, } E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \quad \text{for } r > R. \tag{2.13}$$

Thus for $r > R$, the total distribution of charges may be considered to be concentrated at the centre of the spherical region.

(b) *E for points inside the sphere (r < R)*

Now consider a point P at a distance r from the centre of the sphere of charge. Imagine a spherical Gaussian surface of radius r passing through the point. Again from symmetry consideration, **E** and **dS** will point in the same direction. Applying Gauss' law we obtain

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = \epsilon_0 \oint E\cos 0° dS$$

$$= \epsilon_0 E\oint dS = \epsilon_0 E.4\pi r^2 = \int_V \rho dV = q' \text{ (say)}$$

where q' is the portion of q contained within the sphere of radius r.

$$\therefore \; E = \frac{1}{\epsilon_0}\frac{q'}{4\pi r^2} = \frac{1}{4\pi\epsilon_0}\frac{q'}{r^2}$$

$$\text{Now } q = \rho.\frac{4}{3}\pi R^3 \text{ and } q' = \rho.\frac{4}{3}\pi r^3$$

$$\therefore \; q' = q\frac{\frac{4}{3}\pi r^3}{\frac{4}{3}\pi R^3} = q\left(\frac{r}{R}\right)^3$$

Hence the expression for E for points inside the sphere (r < R) becomes

$$E = \frac{1}{4\pi\epsilon_0} \frac{qr}{R^3} \quad \text{for} \quad r < R. \tag{2.14}$$

Field due to a line of charge

A section of a thin infinitely long wire or a rod (cylindrical conductor) of charge is shown in Fig. 2.7. Let it have a uniform *linear charge density, λ i.e.,* the charge per unit length on the line. We are to find an expression for **E** at a distance r from the line.

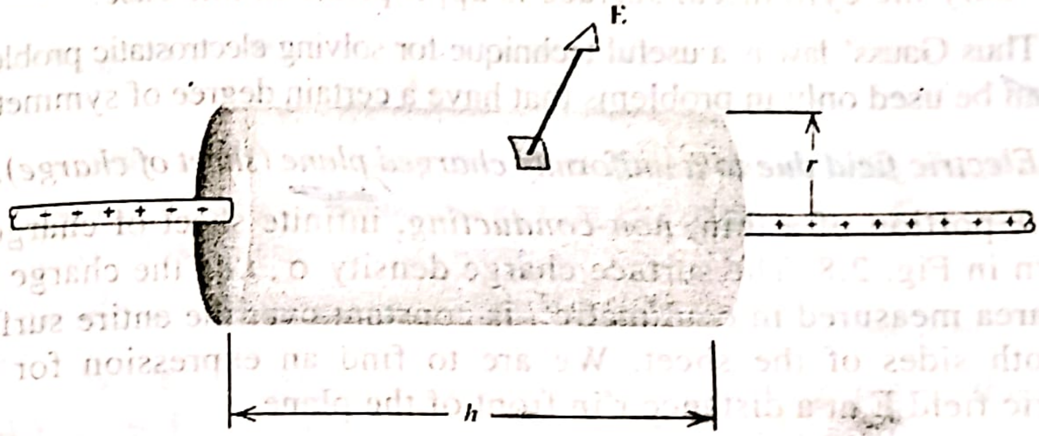


Fig. 2.7

As a Gaussian surface, let us choose a circular cylinder of radius r and length h, closed at each end by plane caps normal to the axis. From symmetry, the field is everywhere constant on the Gaussian surface and for a positive charge directed radially away from the axis so that **E** and **dS** are in the same direction on the curved surface. There will be no flux through the circular caps at two ends because **E** here lies in the surface at every point; therefore the angle between **E** and **dS** is 90° at every point.

Applying Gauss' law, we therefore have,

$$\epsilon_0 \oint \mathbf{E}\cdot d\mathbf{S} = \epsilon_0 \oint E\cos 0^\circ dS$$

$$= \epsilon_0 E\, 2\pi rh = q \tag{2.15}$$

where q is the charge enclosed by the cylinder and $2\pi rh$ is the volume of the cylinder. But $q = \lambda h$. So eqn. 2.15 reduces to

$$\epsilon_0 E\, 2\pi rh = \lambda h$$

$$\text{or, } E = \frac{\lambda}{2\pi\epsilon_0 r} \tag{2.16}$$

The same problem was also solved in Art. 1.3 by using the method of integration. It may be noticed that how much simpler the solution becomes when Gauss' law is applied. It may also be noticed that the solution using Gauss' law is possible only if the Gaussian surface is choosen in such a way that full advantage of the radial symmetry of the electric field set up by a long line of charge can be taken. One is free to choose any surface for a Gaussian surface such as a cube or a sphere. However, they are not all useful for the problem at hand, only the cylindrical surface is appropriate in this case.

Thus Gauss' law is a useful technique for solving electrostatic problems but can be used only in problems that have a certain degree of symmetry.

Electric field due to a uniformly charged plane (sheet of charge).

A portion of a thin, *non-conducting*, infinite sheet of charge is shown in Fig. 2.8. The surface charge density σ, *i.e.*, the charge per unit area measured in coul/metre^2, is constant over the entire surface on both sides of the sheet. We are to find an expression for the electric field **E** at a distance r in front of the plane.

An appropriate Gaussian for the problem is a pill-box of cross-sectional area S and height 2r, arranged to pierce the plane (r on each side) as shown in the figure. From symmetry **E** is uniform at every point and points at right angles to the end caps away from the plane. This means **E** and dS are in the same direction on the end caps. Since **E** does not pierce the cylindrical surface, *i.e.*, **E** and dS are perpendicular to each other at every point on the surface, there is no contribution to the flux

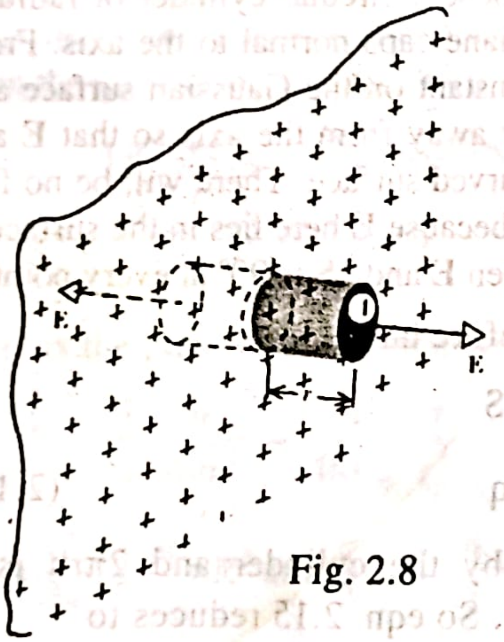


Fig. 2.8

from this source. Hence, applying Gauss' law, we obtain

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = \epsilon_0 \oint E\cos 0° dS = q \qquad (2.17)$$

where q is the charge enclosed by the area of cross-section S.

Now $\oint E\cos 0° dS = (ES + ES)$

And $q = \int_s \sigma \cdot dS = \sigma.S$

So eqn. 2.17 becomes

$$\epsilon_0 (ES + ES) = \sigma S$$

$$\text{or, } 2ES = \frac{\sigma S}{\epsilon_0}$$

$$\text{or, } E = \frac{\sigma}{2\epsilon_0}$$

Eqn. 2.18 shows that the field due to an infinite sheet of charge is independent of the distance from the plane, i.e., E is same for all points on each side of the plane.

(vi) *Electric field due to infinite charged conducting plate*

When a charge is given to a conducting plate, the same is distributed over the entire surface of the plate. If the plate is of uniform *size and thickness*, the surface density (σ) of the plate is uniform and is same on both sides of the surfaces. We are to find an expression for the electric field at a point P in front of the plate.

An appropriate Gaussian surface for the problem is a cylinder through P, which is normal to the surface of the plate. Let S be the cross-sectional area of the cylinder. One end of the cylinder is *inside the conductor* (Fig. 2.9) and so does not contribute to the electric flux as no charge exists inside the conductor. Also no contribution of flux is given by the curved surface of the cylinder as E and dS are perpendicular to each other at every point on this surface. So it is only end cap around P which contributes towards electric flux. So we have

$$\oint \mathbf{E}.d\mathbf{S} = \oint E\cos 0° dS$$
$$= E\oint dS = ES$$

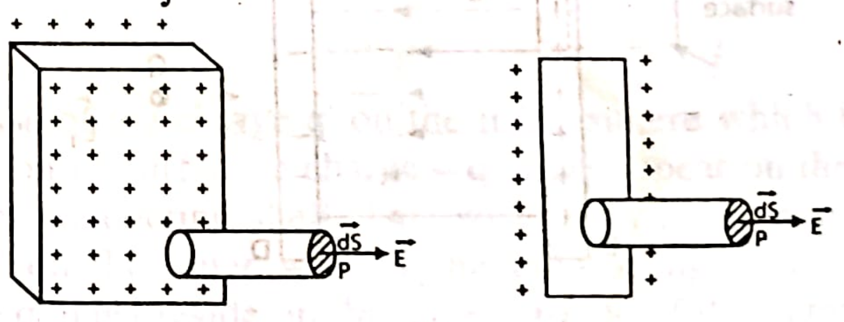


Fig. 2.9

as **E** and **dS** point in the same direction on the end cap.

Charge enclosed by the Gaussian surface of cross-sectional area S is

$$q = \sigma.S.$$

So from Gauss' theorem

$$\therefore \oint \mathbf{E} . d\mathbf{S} = ES = \frac{q}{\epsilon_0} = \frac{\sigma S}{\epsilon_0}$$

$$\text{or, } E = \frac{\sigma}{\epsilon_0} \tag{2.19}$$

It may be noted that the electric field due to a charged conducting plate is twice the field due to a plane sheet of charge having the same surface charge density. It is due to the fact that the same charge is present on its both sides.

(vii) *Electric field due to two parallel charged plates*

Let AB and CD be two charged parallel plates of very great extent. Of the two plates AB has positive charge while CD negative, the charge density σ being same for both the plates.

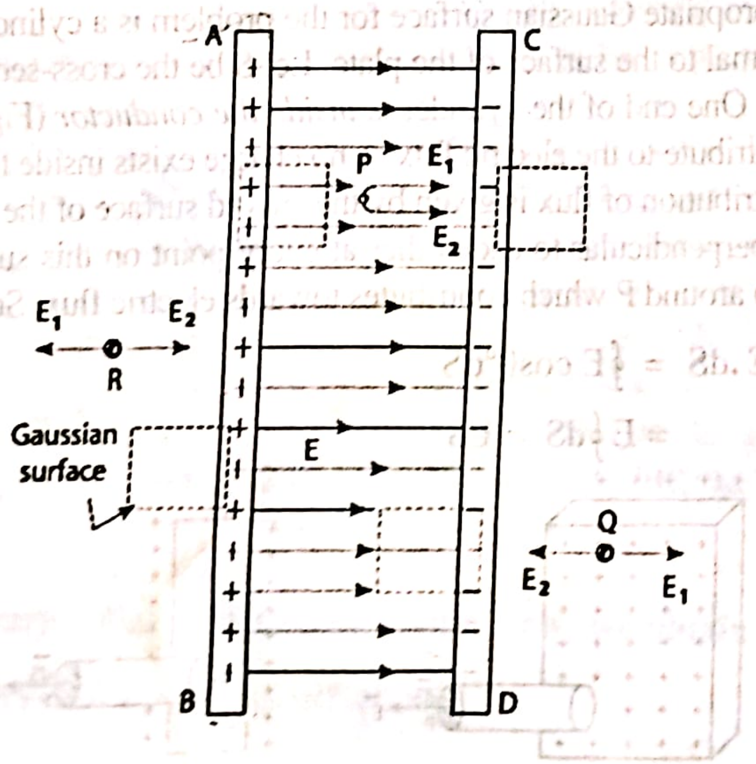


Fig. 2.10

At the point P inside the plates (Fig. 2.10), the field due to AB is $E_1 = \frac{\sigma}{2\epsilon_0}$ (eqn. 2.18) and points to the right. The field due to CD is $E_2 = \frac{\sigma}{2\epsilon_0}$ which also points to the right. [the electric field points away from the positive charge and points towards the negative charge]. Since electric fields superpose, the total field at P is

$$E = E_1 + E_2 = \frac{\sigma}{2\epsilon_0} + \frac{\sigma}{2\epsilon_0} = \frac{\sigma}{\epsilon_0}$$

the direction being from plate AB to the plate CD, *i.e.*, from the positive charge to the negative charge.

Let us now consider a point Q outside the plates. The field due to AB is $\frac{\sigma}{2\epsilon_0}$ towards the right while the field due to CD is also $\frac{\sigma}{2\epsilon_0}$ but points towards the left. Thus the total field at Q is zero. This proves that the field due to parallel charged plates is constant at any point between the plates but vanishes at points outside the plates.

***Example 2.2** A spherical conducting shell of inner radius b and outer radius a carries a charge q. A second spherical conductor of radius c and charge q' is introduced inside it through a hole. Determine the electric fields in the regions (i) outside the outer conductor; (ii) inside the outer spherical shell; (iii) the space between the two spheres; and (iv) within the inner sphere.*

Soln.

Because of the charge q' on the inner sphere which is uniformly distributed on its surface, a charge – q' must appear on the inner walls of the outer conducting shell. This will result in the appearance of a charge +q' on the outer wall of the conducting shell. Therefore a charge q + q' must reside on the outer surface of the outer conducting shell. The arrangement of charges is shown in Fig. 2.11. The electric fields can be obtained by the application of Gauss' law.

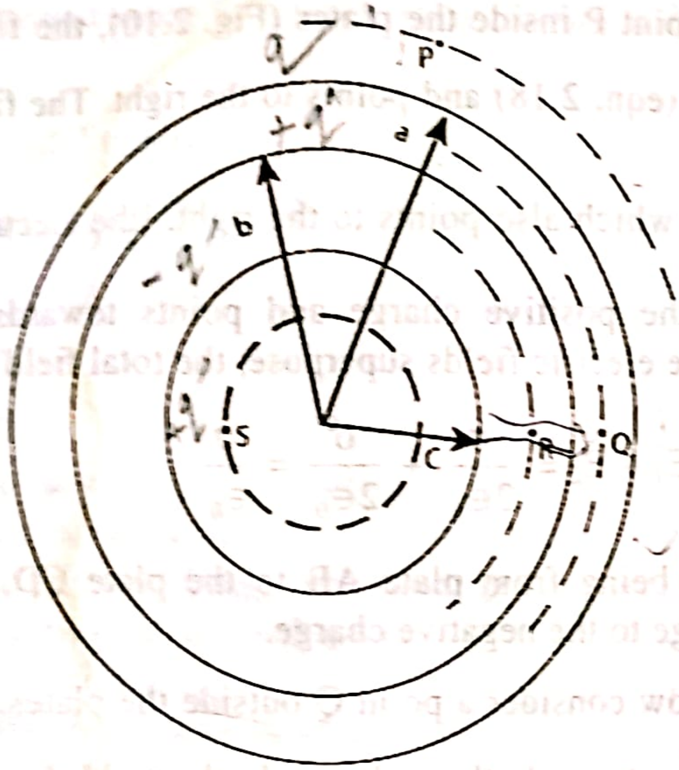


Fig. 2.11

(i) In the region outside the outer conductor, say at point P, we obtain $\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = \sum q$.

Now the total charge enclosed by the Gaussian surface passing through the point P is q + q' + (− q') + q' = q + q'.

$$\therefore\ \epsilon_0 E . 4\pi r^2 = q + q'; \qquad \text{or,}\quad E = \frac{1}{4\pi\epsilon_0}\frac{(q+q')}{r^2}.$$

(ii) At a point inside the spherical shell, say at point Q. The total charge enclosed by the Gaussian surface passing through Q is − q' + q' = 0. Thus, applying Gauss' law we obtain

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = \sum q = 0; \quad \text{or,}\quad \epsilon_0 E 4\pi r^2 = 0; \quad \text{or,}\quad E = 0.$$

(iii) At a point in the space between the two spheres, say at point R. The total charge enclosed by the Gaussian surface passing through R is q'. Hence

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = q'; \quad \text{or,}\quad \epsilon_0 E 4\pi r^2 = q'; \quad \text{or,}\quad E = \frac{1}{4\pi\epsilon_0}\frac{q'}{r^2}.$$

(iv) At a point within the inner sphere, say at a point S. Charge enclosed by the Gaussian surface passing through S is zero.

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{S} = 0; \qquad \text{or,}\quad \epsilon_0 E 4\pi r^2 = 0; \qquad \text{or,}\quad E = 0.$$

Example 2.3 The magnitude of the average electric field normally present in the Earth's atmosphere just above the surface of the Earth is above 150 N/C, directed downward. What is the total net surface charge carried by the Earth? Assume the earth to be a conductor.

Soln.

Lines of force terminate on negative charges so that the Earth's electric field points downward. Hence the average surface charge density σ must be negative. Thus from the relation $E = \sigma / \epsilon_0$ we obtain

$$\sigma = \epsilon_0 E = (8.85 \times 10^{-12}\ C^2/N.m^2)(-150\ N/C)$$
$$= -1.33 \times 10^{-9}\ C/m^2.$$

The Earth's total charge q is given by the product of its surface charge density (σ) and its surface area $4\pi R^2$. Thus

$$q = \sigma.4\pi R^2 = (-1.33 \times 10^{-9}\ C/m^2)(4\pi)(6.37 \times 10^6 m)^2$$
$$= -6.8 \times 10^5\ C$$
$$= -680\ KC.$$

Example 2.4 A plastic rod, whose length L is 220 cm and whose radius R is 3.6 mm, carries a negative charge q of magnitude 3.8×10^{-7}C, spread uniformly over its surface. What is the electric field near the midpoint of the rod, at a point on its surface?

Soln.

Although the rod is not infinitely long, for a point on its surface and near its midpoint, it may be considered to be effectively infinitely long.

The linear charge density for the rod is

$$\lambda = \frac{q}{L} = \frac{-3.8 \times 10^{-7} C}{2.2 m} = -1.73 \times 10^{-7} C/m$$

We then have from eqn. 2.16

$$E = \frac{\lambda}{2\pi\epsilon_0 r} = \frac{-1.73 \times 10^{-7} C/m}{(2\pi)(8.85 \times 10^{-12} C^2/N.m^2)(0.0036m)}$$
$$= -8.6 \times 10^5\ N/C.$$

***Example 2.5** Portions of two large sheets of charge with uniform surface charge densities s+ = + 6.8 mC/m² and s− = − 4.3mC/m² are shown in Fig. 2.12. Find the electric field E (a) to the left of the sheets, (b) between the sheets, and (c) to the right of the sheets.*

Soln.

The electric field due to each sheet separately will be first determined. The resultant field will then be determined by using the principle of superposition.

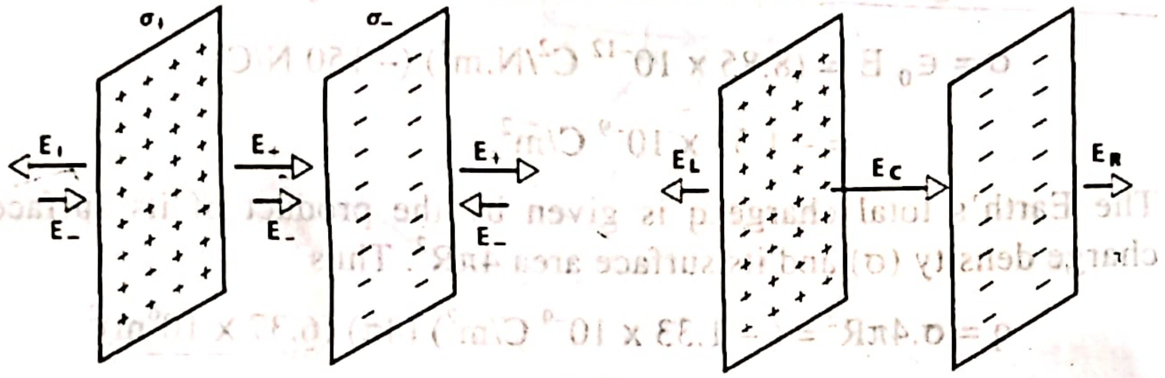


Fig. 2.12

For the electric field due to the positive sheet we have,

$$E_+ = \frac{\sigma_+}{2\epsilon_0} = \frac{6.8\times10^{-6}C/m^2}{(2)(8.85\times10^{-12}C^2/N.m^2)} = 3.84 \times 10^5 N/c.$$

Similarly for the negative sheet,

$$E_- = \frac{\sigma_-}{2\epsilon_0} = \frac{4.3\times10^{-6}C/m^2}{(2)(8.85\times10^{-12}C^2/N.m^2)} = 2.43 \times 10^5 N/c.$$

The direction of the fields are shown in the figure for left of the sheets, between the sheets and right of the sheets. The resultant fields in these regions are obtained following the principle of superposition. For a point on the left of the sheets we have,

$$E_L = -E_+ + E_- = -3.8 \times 10^5 \text{ N/C} + 2.43 \times 10^5 \text{N/C}$$

$$= -1.4 \times 10^5 \text{ N/C}.$$

The resultant field therefore points towards the left.

For a point on the right, the electric field E_R has the same magnitude but points towards the right.

110211

Between the sheets, the two fields point to the same direction. They, therefore, add to give

$$E_C = 3.84 \times 10^5 \text{ N/C} + 2.43 \times 10^5 \text{ N/C} = 6.3 \times 10^5 \text{ N/C}.$$

Example 2.6 Two parallel plates of area 20 cm² each are separated by a distance of 10 mm. If one plate carries a charge +2 x 10⁻⁹C, and the other carries a charge –2 x 10⁹C, calculate the electric field in the space between them.

Soln.

Surface charge density for each of the plate is given by

$$\sigma = \frac{2\times10^{-9}}{20\times10^{-4}} \text{Cm}^{-2} = 10^{-6} \text{ Cm}^{-2}.$$

Since one of the plate carries positive charge while the other carries negative charge, the electric field in the space between them

$$E = \frac{\sigma}{\epsilon_0} = \frac{10^{-6}\text{Cm}^{-2}}{8.85\times10^{-12}\text{C}^2\text{N}^{-1}\text{m}^{-2}} = 1.13 \times 10^5 \text{ N/C}.$$

Example 2.7 The electric strength of air is about 3.0 x 10⁶ N/C which means that if the electric field exceeds this value, sparking will occur. What is the largest charge a 0.50 cm. radius sphere can hold if sparking is not to occur in the air surrounding it?

Soln.

For a point outside a sphere, all the charge on the surface can be considered to be located at the centre of the sphere. Thus the value of q can be obtained from the relation.

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$$

$$\text{or, } 3.0 \times 10^6 \text{ N/C} = \frac{1}{(4\pi)(8.85\times10^{-12}\text{C}^2/\text{N.m}^2)}\frac{q}{(5.0\times10^{-3}\text{m})^2}$$

$$\text{or, } q = 8.34 \times 10^{-9}\text{C} = 8.34\text{nC}.$$

EXERCISES

1. Define electric flux. Give an expression for the flux in an electric field.

2. State and prove Gauss' theorem and show that Coulomb's law of electrostatics can be deduced from it.

3. State Gauss' theorem and apply it to obtain an expression for the electric field at a point (i) outside and (ii) inside a charged conducting sphere.

4. Prove that while determining the electric field at a point outside a charged conducting sphere, the entire charge can be considered to be located at the centre of the sphere.

5. State and prove Gauss' law of electrostatics. Apply it to find the electric field due to a line of charge (cylindrical conductor).

6. Find an expression for the electric field at a point due to a plane sheet of charge. Hence prove that the electric field is independent of the distance of the point.

7. Show that the electric field at a point between two similarly but oppositely charged plates is constant but zero outside the plates.

8. The maximum electric intensity which air can withstand at atmospheric pressure is about 3×10^6 N/C. Calculate the maximum charge density which a metallic conductor of radius R metres can have when placed in air. [26×10^{-6} Cm^{-2}]

9. How much charge must be added to an isolated spherical conductor of 8 cm radius to produce a field just outside the surface whose intensity is 1000 N/C?

10. Calculate the electric field at the surface of a gold nucleus (Z = 79, $r = 6.9 \times 10^{-15}$ m, $e = 1.6 \times 10^{-19}$ C)

11. The plates of a capacitor carry charge +Q and –Q as shown. Each plate has area A = 600 cm^2. Between the plates, the field is constant at E = 300 KV/m and the field is zero outside the plates. Evaluate Q. [0.159μC]

12. A metal can suspended by a silk thread carries a positive charge (not uniformly distributed). At a certain point very near the surface of the can, the electric field has a value E = 600 KV/m. Evaluate the surface charge density σ near the point. [5.313 $\mu C/m^2$]

CHAPTER III

ELECTRIC POTENTIAL

3.1 Electric Potential

The electric field around a charged particle can be described not only by the electric field strength E which is a vector quantity but also by a scalar quantity – the *electric potential* V. The two quantities are intimately related, and often it is a matter of convenience which is used in a given problem.

The idea of electric potential is related to work done in carrying a charge from one point to another in an electric field. Let us begin by placing a free positive charge q_0 in an electrostatic field. It will experience a force in the direction of the electric intensity. A free negative charge will move in the opposite direction. Work has to be done in moving the free positive charge – referred to as the *test charge*, against the direction of the field. The work done in moving a test charge q_0 from a point A to a point B with constant speed, *i.e.*, always keeping the charge in *equilibrium*, is called the *electric potential difference* between the points. If the electric potential difference between the points is $V_B - V_A$, then it is defined by the equation

$$V_B - V_A = \frac{W_{AB}}{q_0} \tag{3.1}$$

where W_{AB} is the work done by an agent to move the charge from A to B. The work W_{AB} may be (i) positive, (ii) negative, or (iii) zero. In these cases the electric potential at B will be (i) higher, (ii) lower, or (iii) the same as the electric potential at A.

The mks unit of potential difference that follows from eqn. (3.1) is joule/coulomb. However, this combination occurs so often that a special unit, the *volt*, is used to represent it. Or,

1 volt = 1 joule/coul

The potential difference between two points is 1 volt if one joule of work is done in moving one coulomb of charge from one point to the other.

W_{AB} is also the increase in potential energy because there is no increase in kinetic energy as the test charge is made to move at constant speed. Thus we expect the potential to rise for displacements in opposition to **E** and to fall for displacements along **E**.

Strictly speaking, only differences of potential can be defined. We can, however, always choose the point A to be at large (strictly an infinite) distance from all charges, and the electric potential V_A at this infinite distance is arbitrarily taken as zero. Putting $V_A = 0$ in eqn. (3.1) and dropping the subscript leads to

$$V = \frac{W}{q_0} \tag{3.2}$$

where W is the work that an external agent must do to move that test charge q_0 from infinity to the point in question. Thus the *electric potential at any point is defined as the work that must be done in bringing a unit positive charge from infinity up to the point.* Since both W and q_0 in eqn. (3.2) are scalar quantities, electric potential is a scalar quantity.

Bearing in mind the assumptions made about the reference position, electric potential near an isolated positive charge is positive since positive work must be done by the outside agent to push a (positive) test charge in from infinity. Similarly, the potential near an isolated negative charge is negative because an outside agent must exert a restraining force *i.e.*, must do negative work on positive test charge as it comes in from unity.

It can be easily proved that both W_{AB} and $V_B - V_A$ are independent of the path followed in moving the test charge from point A to point B. If this were not so, point B would not have a unique electric potential (with respect to point A as a defined reference position) and the concept of potential would have limited usefulness.

3.2 Potential and Electric Field Strength

Fig. 3.1 shows two points A and B in a uniform electric field **E**, set up by an arrangement of charge not shown in the figure. Let the distance of A from B in the field direction be d. Assume that a

positive test charge q_0 is being moved by an external agent from A to B along the straight line connecting them under the condition of equilibrium *i.e.*, q_0 moves in such a way that it is not accelerated; it moves with a constant velocity.

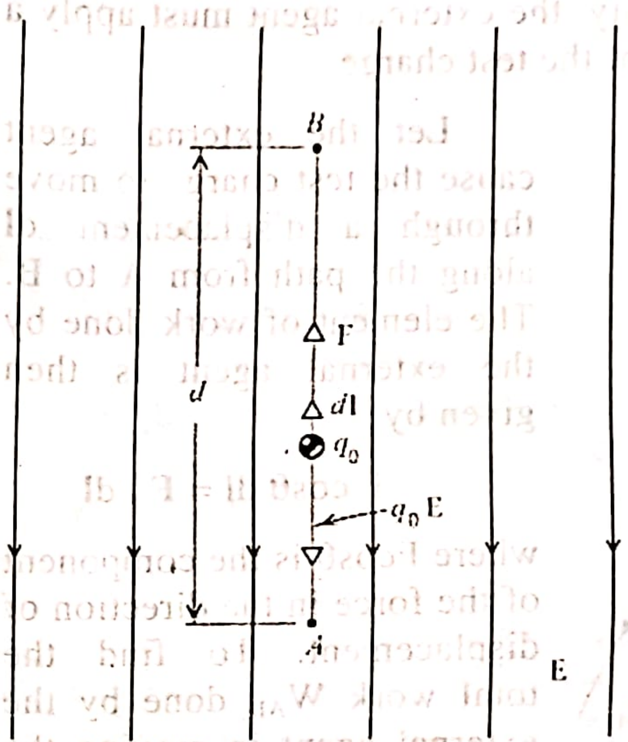


Fig. 3.1

The electric force on the charge is q_0E and points downward. To move the charge in the manner described above, the force on q_0 must be countered by an external force F of the same magnitude but directed upward. The work done by the agent that supplies this force is

$$W_{AB} = F.d = q_0Ed \qquad (3.3)$$

Substituting this in eqn. 3.1, we have

$$V_B - V_A = \frac{W_{AB}}{q_0} = \frac{q_0Ed}{q_0} = Ed \qquad (3.4)$$

Eqn. (3.4) gives the relation connecting potential difference and field strength for a simple special case. From eqn. (3.4) it appears that another unit for E $\left(=\frac{V_B - V_A}{d}\right)$ is *volt per metre*. It can be proved that volt/metre is identical with newton/coulomb. Or,

1 volt/metre = 1 nt/coulomb.

In Fig. 3.1 B is at a higher potential than A. This is expected because the external agent would have to do positive work to push a positive charge from A to B against the direction of the field.

Let us now investigate the relation between V and E in the more general case in which the field is *not* uniform and in which the test body is made to move along a path which is *not* straight. The electric field exerts a force q_0E on the test charge as shown in Fig. 3.2. To keep the test charge from accelerating *i.e.*, if the test charge is to move with a constant velocity, the external agent must apply a force $\mathbf{F} = -q_0\mathbf{E}$ for all positions of the test charge.

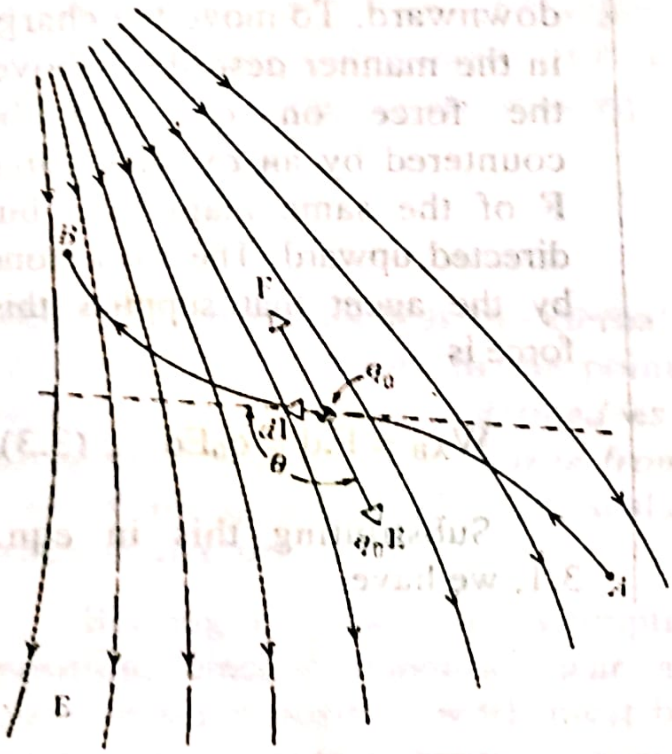


Fig. 3.2

Let the external agent cause the test charge to move through a displacement dl along the path from A to B. The element of work done by the external agent is then given by

$$F\cos\theta\, dl = \mathbf{F}\cdot\mathbf{dl}$$

where $F\cos\theta$ is the component of the force in the direction of displacement. To find the total work W_{AB} done by the external agent in moving the test charge from A to B, we integrate (add up) the work contributions for all infinitesimal segments into which the path is divided. This leads to

$$W_{AB} = \int_A^B \mathbf{F}\cdot\mathbf{dl} = -q_0\int_A^B \mathbf{E}\cdot\mathbf{dl} \qquad (3.5)$$

since $\mathbf{F} = -q_0\mathbf{E}$

The type of integral as given by eqn. (3.5) is called a *line integral.*

Substituting this expression for W_{AB} in eqn. (3.1), we obtain

$$V_B - V_A = \frac{W_{AB}}{q_0} = \frac{-q_0\int_A^B \mathbf{E}\cdot\mathbf{dl}}{q_0}$$

$$= -\int_A^B \mathbf{E} . d\mathbf{l} \tag{3.6}$$

If the point A is at infinite distance then the potential V_A at infinity is taken as zero. Eqn. (3.6) then gives the potential V at the point B, or, dropping the subscript B,

$$V = -\int_{\infty}^{B} \mathbf{E} . d\mathbf{l} \tag{3.7}$$

Eqns. (3.6) and (3.7) allow us to calculate either the potential difference between any two points or the potential at any point if **E** is known at various points in the field.

Example 3.1 Calculate $V_B - V_A$ in Fig. 3.1 using eqn. (3.6). Compare the result with that obtained by direct analysis of this special case by using eqn. (3.4)

Soln.

In moving the test charge from A to B the path dl always points upward while the electric field **E** points downward (Fig. 3.1), so that the angle θ between **E** and **dl** is 180°.

Eqn. (3.6) then becomes

$$V_B - V_A = -\int_A^B \mathbf{E} . d\mathbf{l} = -\int_A^B E \cos 180^\circ \, dl = \int_A^B E dl$$

E is constant for all parts of the path in this problem and hence can be taken outside the integral sign, giving

$$V_B - V_A = E \int_A^B dl = Ed$$

the same as that given by eqn. (3.4).

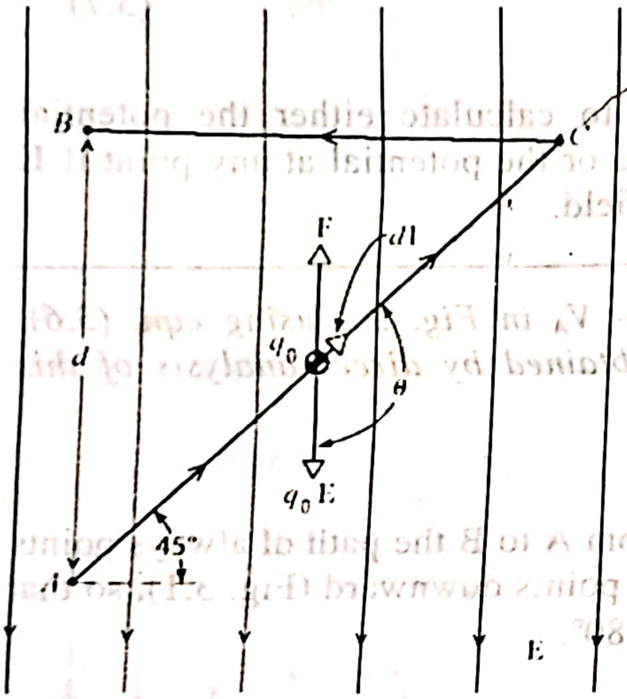
Example 3.2 Let a test charge q_0 be moved without acceleration from A to B over the path as shown in Fig. 3.3 Compute the potential difference between A and B.

Soln.

For the path AC, θ = 135° and, from eqn. (3.6), we obtain

$$V_C - V_A = -\int_A^C \mathbf{E} \cdot d\mathbf{l} = -\int_A^C E \cos 135° dl$$

$$= \frac{E}{\sqrt{2}} \int_A^C dl$$



The integral is the length of the line AC which is $\sqrt{2}$ d. Thus

$$V_C - V_A = \frac{E}{\sqrt{2}}(\sqrt{2}\, d) = Ed$$

Points B and C are at the same potential because no work is done in moving a charge between them, **E** and **dl** being at right angles for all points on the line CB. In other words B and C lie on the same equipotential surface at right angles to the lines of force. Thus

$$V_B - V_A = V_C - V_A = Ed$$

This is the same value derived for a direct path connecting A and B. This is to be expected because the potential difference between two points is path independent.

Example 3.3 The electric field outside a long charged wire is given by E = –5000/r V/m and is radially inward. What is the sign of the charge on the wire? Find the value of $V_B - V_A$ if r_B = 60 cm and r_A = 30cm. Which point is at the higher potential?

Soln.

Since the field is directed toward the wire, the wire is charged negatively. Going from A to B is opposite to the direction of the field, so B is at a higher potential than A. Thus

$$V_B - V_A = -\int_{0.3}^{0.6} \frac{-5000}{r}\, dr = 5000\, ln2 = 3470 \text{ V.}$$

3.3 Potential due to a point charge

As shown in Fig. 3.3, A and B are two points near an isolated point charge q. For simplicity we assume that A, B and q lie on a straight line. We shall calculate the potential difference between points A and B, assuming that a test charge q_0 is moved without acceleration along a radial line from A to B.

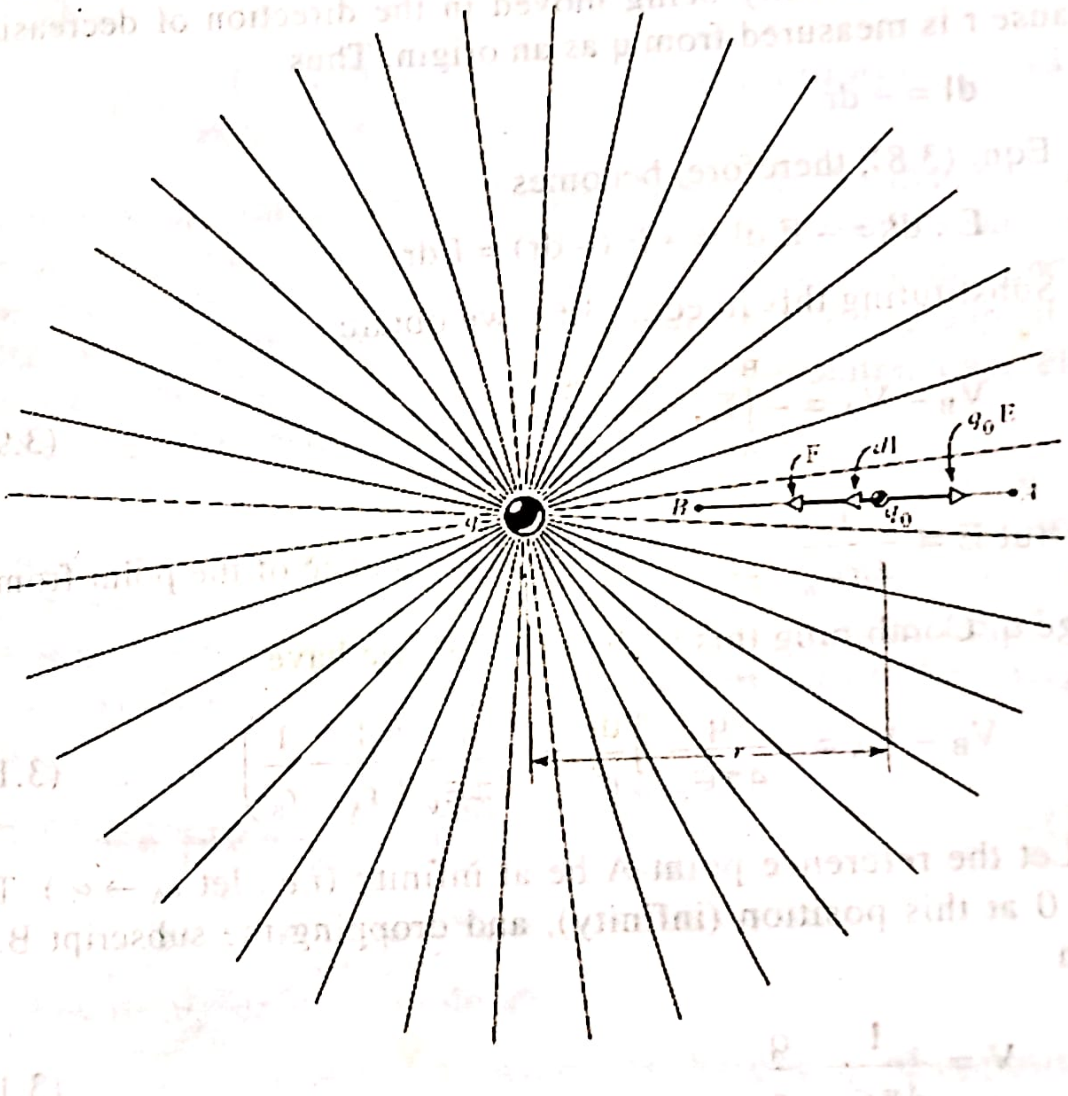


Fig. 3.3

The potential difference between the points A and B is given by (eqn. 3.6),

$$V_B - V_A = -\int_A^B \mathbf{E}.d\mathbf{l}$$

As can be seen in the figure, **E** points to the right whereas dl, which is always in the direction of motion, points to the left; so that the angle between them is $\theta = 180°$. Therefore,

$$\mathbf{E} \cdot d\mathbf{l} = E \cos\theta \, dl = -Edl \quad (3.8)$$

However, as the test charge is moved through a distance dl to the left, it is actually being moved in the direction of decreasing r because r is measured from q as an origin. Thus

$$dl = -dr$$

Eqn. (3.8), therefore, becomes

$$\mathbf{E} \cdot d\mathbf{l} = -E\,dl = -E(-dr) = Edr$$

Substituting this in eqn. (3.6), we obtain

$$V_B - V_A = -\int_A^B \mathbf{E}.d\mathbf{l} = -\int_{r_A}^{r_B} Edr \quad (3.9)$$

But $E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$ where r is the distance of the point from the charge q. Combining this with eqn. (3.9) we have

$$V_B - V_A = -\frac{q}{4\pi\epsilon_0}\int_{r_A}^{r_B}\frac{dr}{r^2} = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r_A} - \frac{1}{r_B}\right) \quad (3.10)$$

Let the reference point A be at infinity (*i.e.*, let $r_A \to \infty$). Then $V_A = 0$ at this position (infinity), and dropping the subscript B, we obtain

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r} \quad (3.11)$$

Eqn. (3.11) clearly shows that *equipotential surfaces for an isolated point charge are spheres concentric with the point charge.*

3.4 Potential due to collection of charges

For a collection of discrete charges, the potential at a point is calculated due to each individual charge, as if the other charges were not present. These potentials are then added and we obtain

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r_1} + \frac{1}{4\pi\epsilon_0}\frac{q_2}{r_2} + \ldots\ldots\ldots \text{ (zero at infinity)}$$

where r_1 is the distance of the point from the charge q_1, r_2 from q_2, etc.

Or,

$$V = \frac{1}{4\pi\epsilon_0}\sum_n \frac{q_n}{r_n} \quad \text{(zero at infinity)} \qquad (3.12)$$

where q_n is the value of the n^{th} charge and r_n is the distance of this charge from the point in question. Since potentials are scalar quantities, the sum used to calculate V is an *algebraic sum* and not a vector sum like the one used to calculate **E** for a group of point charges. This is an important advantage of potential over electric field strength for making calculations.

If the charge distribution is continuous, instead of being discrete, then eqn. (3.11) can be used to find dV at a point due to a typical element of charge dq. The total potential V at the point is then obtained by integrating over the whole region occupied by the charge. Thus, if r is the distance of the point from the charge dq, then

$$V = \int dV = \frac{1}{4\pi\epsilon_0}\int\frac{dq}{r} \qquad (3.13)$$

2015-16 2013-14

3.5 Potential due to a dipole

As discussed in Art. 1.5, two equal charges, q, of opposite sign, separated by a small distance 2a, constitute an electric dipole. The electric dipole moment **p** has the magnitude 2aq and points from the negative charge to the positive charge. We would like to derive an expression for the electric potential V at any point of space due to a dipole, provided only that the point is not too close to the dipole.

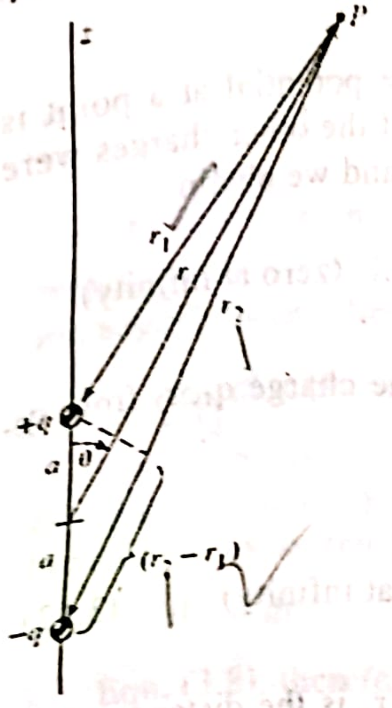


Fig. 3.4

Let the point P, where we would like to calculate the potential, be specified by the quantities r and θ in Fig. 3.4. From symmetry, it is clear that the potential will not change as point P rotates about the z axis, r and θ being fixed. Thus we need only find V (r, θ) for any plane containing this axis; the plane of Fig. 3.4. being one such plane.

Let us first calculate the resultant potential at the point P due to charges +q and −q. From eqn. 3.11 we have

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{q}{r_2} - \frac{1}{4\pi\epsilon_0}\frac{q}{r_1}$$

$$= \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r_2} - \frac{q}{r_1}\right) = \frac{q}{4\pi\epsilon_0}\left(\frac{r_1 - r_2}{r_1 r_2}\right) \qquad (3.14)$$

But $r_1^2 = r_2^2 + (2a)^2 + 2r_2\,(2a)\cos\theta'$

or, $r_1^2 - r_2^2 = 2a\,(2a + 2r_2\cos\theta')$

$$\text{or, } r_1 - r_2 = \frac{2a(2a + 2r_2\cos\theta')}{r_1 + r_2} \qquad (3.15)$$

Substituting the value of $r_1 - r_2$ as given by eqn. (3.15) in eqn.(3.14) we obtain

$$V_P = \frac{q.2a(2a + 2r_2\cos\theta')}{4\pi\epsilon_0\, r_1 r_2 (r_1 + r_2)} \qquad (3.16)$$

Eqn. (3.16) gives the exact value of the potential at the point P. But for an ideal dipole ($2a \ll r$) $r_1 \rightarrow r$, $r_2 \rightarrow r$ and $\theta' \rightarrow \theta$, and in the limit

$$V_P = \frac{q.2a\cos\theta}{4\pi\epsilon_0 r^2} \tag{3.17}$$

The quantity q.2a is the electric dipole moment p of the dipole. Hence eqn. (3.17) reduces to

$$V_P = \frac{p\cos\theta}{4\pi\epsilon_0 r^2} \tag{3.18}$$

Special cases

(i) when the point P lies on the axial line of the dipole on the side of the positive charge, $\theta = 0$, $\therefore \cos\theta = 1$.

So from eqn. (3.18)

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{p}{r^2} \tag{3.19}$$

(ii) when the point P lies on the axial line of the dipole but on the side of the negative charge, $\theta = 180°$, $\therefore \cos\theta = -1$.

$$\therefore\ V_P = -\frac{1}{4\pi\epsilon_0}\frac{p}{r^2} \tag{3.20}$$

(iii) when the point P lies on the equatorial line of the dipole, $\theta = 90°$, $\therefore \cos\theta = 0$.

$$\therefore\ V = 0 \tag{21}$$

Eqn. (3.21) reflects the fact that *no work is done in bringing a charge from infinity to the dipole along the perpendicular bisector of the dipole.*

Radial and transverse components of the electric vector (E_r and E_θ)

The radial component of the electric vector E_r (in the direction OP) is obtained from the relation

$$E_r = -\frac{\partial V}{\partial r} = -\frac{d}{dr}\left(\frac{p\cos\theta}{4\pi\epsilon_0 r^2}\right)$$

$$= \frac{2p\cos\theta}{4\pi\epsilon_0 r^3} = \frac{p\cos\theta}{2\pi\epsilon_0 r^3} \quad (3.22)$$

The transverse component E_θ in the direction perpendicular to OP is given by

$$E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta}$$

$$= -\frac{1}{r}\frac{d}{d\theta}\left(\frac{p\cos\theta}{4\pi\epsilon_0 r^2}\right)$$

$$= \frac{p\sin\theta}{4\pi\epsilon_0 r^3} \quad (3.23)$$

The resultant field intensity is given by

$$E^2 = E_r^2 + E_\theta^2$$

$$= \left[\left(\frac{2p\cos\theta}{4\pi\epsilon_0 r^3}\right)^2 + \left(\frac{p\sin\theta}{4\pi\epsilon_0 r^3}\right)^2\right]$$

$$= \frac{p^2}{16\pi^2\epsilon_0^2 r^6}(1 + 3\cos^2\theta)$$

$$\therefore\ E = \frac{p}{4\pi\epsilon_0 r^3}(1 + 3\cos^2\theta)^{1/2} \quad (3.24)$$

The resultant field intensity is inclined at an angle φ to r such that

$$\tan\varphi = \frac{p\sin\theta}{4\pi\epsilon_0 r^3} \times \frac{4\pi\epsilon_0 r^3}{2p\cos\theta} = \frac{1}{2}\tan\theta$$

When $\theta = 0$, $V = \frac{p}{4\pi\epsilon_0 r^2}$ and $E_r = \frac{p}{4\pi\epsilon_0 r^3}$ along the axis.

When $\theta = \frac{\pi}{2}$, $V = 0$ and $E_\theta = \frac{p\sin\theta}{4\pi\epsilon_0 r^3}$ on the right bisector and perpendicular to the axis and $E_r = 0$ for all values of r. Hence the right bisector of the dipole is the equipotential plane for which $V = 0$.

3.6 Potential due to an electric quadrupole

An electric quadrupole consists of two electric dipoles placed end to end along the same line. Fig. 3.5 shows one such arrangement where OA and OB are two dipoles each of length a. The quadrupole is so arranged that they almost, but not quite cancel each other in their electric effects at distant points. Let us calculate the electric potential at a point P along the axis of the quadrupole.

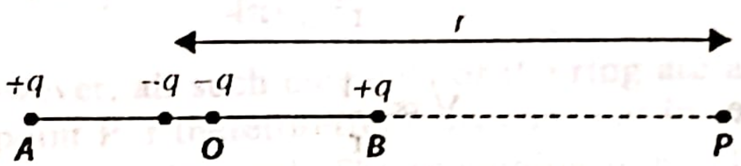


Fig. 3.5

It is clear from the figure that the potential at the point P is given by

$$V = \frac{1}{4\pi\epsilon_0}\left[\frac{q}{(r-a)} - \frac{q}{r} - \frac{q}{r} + \frac{q}{(r+a)}\right]$$

$$= \frac{1}{4\pi\epsilon_0}\left[\frac{r(r+a) - 2(r^2 - a^2) + r(r-a)}{r(r^2 - a^2)}\right]$$

$$= \frac{q}{4\pi\epsilon_0}\left[\frac{r^2 + ra - 2r^2 + 2a^2 + r^2 - ra}{r(r^2 - a^2)}\right]$$

$$= \frac{1}{4\pi\epsilon_0}\frac{2q.a^2}{r(r^2 - a^2)}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{Q}{r(r^2 - a^2)} = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^3\left(1 - \frac{a^2}{r^2}\right)} \qquad (3.25)$$

where Q (= $2qa^2$) is the *quadrupole moment* of the charge assembly of Fig. 3.5.

If r >> a, $\frac{a}{r} << 1$. Hence $\frac{a^2}{r^2}$ can be neglected compared to 1.

$$\therefore \quad V = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^3}$$

So in the special case of

(i) monopole (single charge) $V \propto \frac{1}{r}$

(ii) dipole $V \propto \frac{1}{r^2}$

(iii) quadrupole $V \propto \frac{1}{r^3}$

3.7 Potential due to ring of charge

Fig. 3.6 shows a uniformly charged thin ring of radius R carrying a uniform linear charge density λ around its circumference. We would like to calculate the potential at a point P, a distance z from the plane of the ring along its central axis. The same method which was followed to calculate the electric field of a continuous charge distribution will also be followed here. The calculation is simpler in the case of the potential, because the potential being a scalar, it is not necessary to take into account the different directions of the contributions from the different elements of charge.

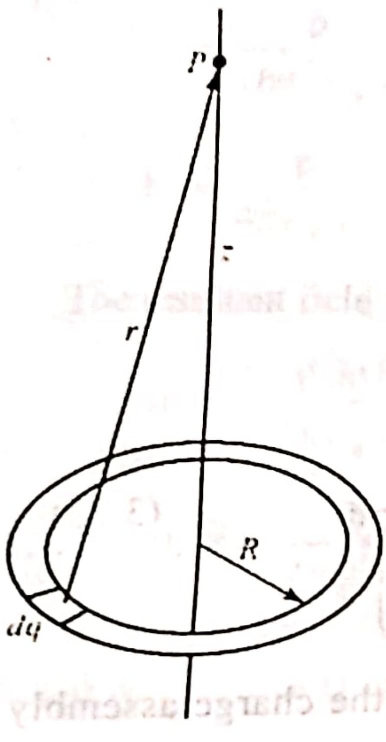


Fig. 3.6

Let us consider a differential element of the ring of length ds located at an arbitrary position on the ring. It contains an element of charge dq = λ.ds.

Each element dq can be considered as a point charge. The contribution dV to the potential at the point P due to this charge is given by

$$dV = \frac{1}{4\pi\epsilon_0}\frac{dq}{r}$$

where r is the distance of the element dq from P.

To find the contribution due to the entire distribution, it is necessary to integrate the individual contributions of all the elements, or

$$V = \int dV = \frac{1}{4\pi\epsilon_0}\int\frac{dq}{r} \tag{3.26}$$

However, all such elements of the ring are at the same distance r from point P. r therefore remains constant in eqn. 3.26 and can be taken out of the integral. The remaining integral $\int dq$, gives simply the total charge q on the ring. The potential at P can therefore be expressed as

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}$$

But $r = \sqrt{R^2 + z^2}$; so

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{\sqrt{R^2 + z^2}} \tag{3.27}$$

3.8 Potential due to a charged disk

Fig. 3.7 shows a uniformly charged circular disk of radius *a* whose surface charge density is σ. We would like to calculate the electric potential at a point P on the axis of the disk at a distance r from the plane of the disk.

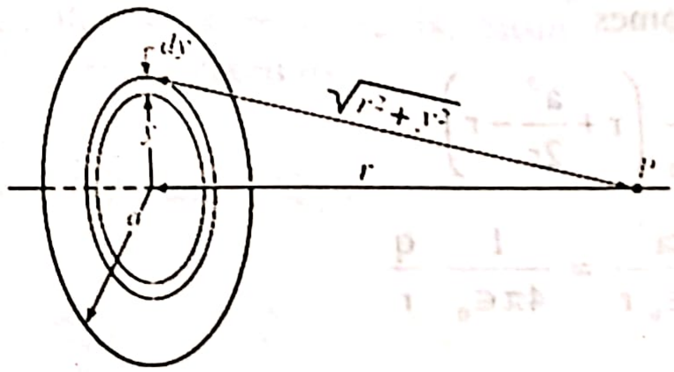


Fig. 3.7

Let us consider a charge element dq consisting of a flat circular strip of radius y and width dy. Then

$$dq = \sigma \cdot (2\pi y)\,(dy)$$

where $(2\pi y)(dy)$ is the area of the strip. All parts of this charge element are at the same distance r' $(= \sqrt{y^2 + r^2})$ from the point P so that their contribution dV to the electric potential at P is given by

$$dV = \frac{1}{4\pi\epsilon_0}\frac{dq}{r'} = \frac{1}{4\pi\epsilon_0}\frac{\sigma.2\pi y dy}{\sqrt{y^2+r^2}}$$

The potential V is found by integrating over all the strips into which the disk can be divided.

Or,

$$V = \int dV = \frac{\sigma}{2\epsilon_0}\int_0^a \left(y^2+r^2\right)^{-\frac{1}{2}} y\, dy$$

$$= \frac{\sigma}{2\epsilon_0}\left(\sqrt{a^2+r^2} - r\right)$$

Eqn. 3.28 is valid for all values of r. If r >> a, the quantity $\sqrt{a^2+r^2}$ can be expanded by binomial theorem as

$$\sqrt{a^2+r^2} = r\left(1+\frac{a^2}{r^2}\right)^{\frac{1}{2}} = r\left(1+\frac{1}{2}\frac{a^2}{r^2}+\ldots\ldots\right)$$

$$\cong r + \frac{a^2}{2r} \text{ (neglecting higher power of } \frac{a^2}{r^2})$$

V then becomes

$$V \cong \frac{\sigma}{2\epsilon_0}\left(r + \frac{a^2}{2r} - r\right)$$

$$= \frac{\sigma\pi a^2}{4\pi\epsilon_0 r} = \frac{1}{4\pi\epsilon_0}\frac{q}{r} \qquad (3.29)$$

where q ($= \sigma\pi a^2$) is the total charge on the disk. This limiting result is expected because for r >> a the disk behaves like a point charge.

3.9 Electric potential energy

Fig. 3.8 shows two charges q_1 and q_2 a distance r apart. If the separation between them is increased, work must be done by an external agent. The work will be positive if the charge are opposite in sign and negative otherwise. The energy represented by this work can be thought of as stored in the system q_1+q_2 as *electric potential energy*. Like all forms of potential energy, this energy can also be transformed into other forms. For example, if q_1 and q_2 are charges of opposite sign and we release them, they will accelerate towards each other, transforming the stored potential energy into kinetic energy of the accelerating masses.

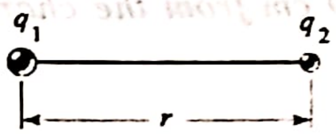


Fig. 3.8

The electric potential energy of a system of point charges may be defined as the work required to assemble this system of charges by bringing them in from an infinite distance. We assume that the charges are all at rest when they are infinitely separated, that is, they have no initial kinetic energy.

Let us imagine that q_2 in Fig. 3.8 to be removed to infinity and at rest. The *electric potential* at the original site of q_2 due to q_1 is given by

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r} \tag{3.30}$$

If q_2 is now moved from infinity to its original position at a distance r from q_1, the work required is, from the definition of electric potential (V = W/q), given by

$$W = V.\, q_2 \tag{3.31}$$

Combining eqns. 3.30 and 3.31

we obtain

$$W = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r} \tag{3.32}$$

But the work is precisely the *electric potential energy* U of the system $q_1 + q_2$. Thus

$$U\,(=W) = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r_{12}}$$

The subscript of r emphasizes that the distance involved is that between the point charges q_1 and q_2.

For a system containing more than two charges, the procedure is to compute the potential energy for every pair of charge separately. The results are then added separately. This procedure rests on a physical picture in which (i) charge q_1 is brought into position, (ii) q_2 is brought from infinity to its position near q_1, (iii) q_3 is brought from infinity to its position near q_1 and q_2, etc.

Example 3.4 What must the magnitude of an isolated positive point charge be for the electric potential at 10 cm from the charge to be +100 volts?

Soln.

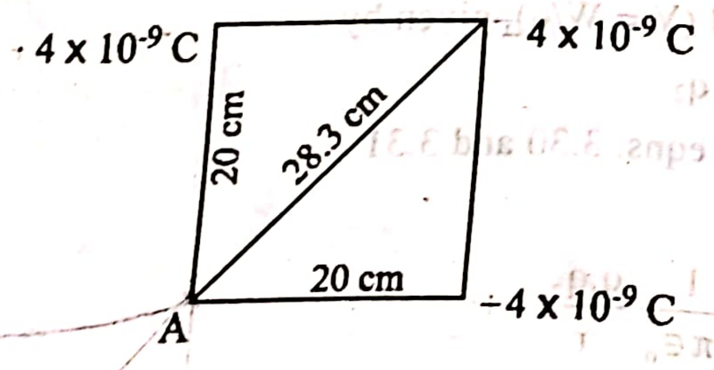
From $V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}$ we obtain

$$q = (V)\,(4\pi\epsilon_0)\,(r)$$

$$= (100 \text{ volts})\,(4\pi)\,(8.9 \times 10^{-12} \text{ coul}^2/\text{nt-m}^2)\,(0.10 \text{ m})$$

$$= 1.1 \times 10^{-9} \text{ coul.}$$

Example 3.5 Three charges are placed at three corners of a square as shown in the figure. Find the potential at point A.



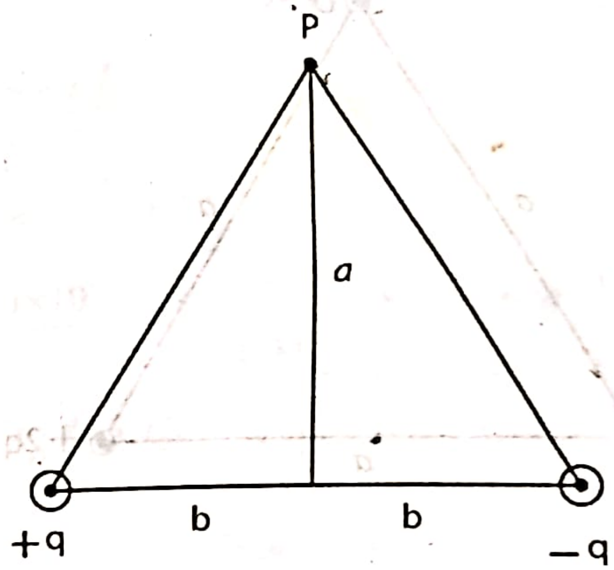
Soln.

$$V_A = \frac{1}{4\pi\epsilon_0}\left[\frac{+4\times10^{-9}C}{0.20m}+\frac{-4\times10^{-9}C}{0.283m}+\frac{+4\times10^{-9}C}{0.20m}\right]$$

$$= (9 \times 10^9)\left[\frac{4\times10^{-9}C}{0.20m}+\frac{-4\times10^{-9}C}{0.283m}+\frac{4\times10^{-9}C}{0.20m}\right]$$

$= 233$ V.

Example 3.6 Show that the absolute potential at point P in the figure below is zero.



Soln.

$$V_P = \frac{1}{4\pi\epsilon_0}\frac{+q}{\sqrt{a^2+b^2}}+\frac{1}{4\pi\epsilon_0}\frac{-q}{\sqrt{a^2+b^2}}$$

$= 0.$

Example 3.7 Two protons in a nucleus of U^{238} are 6.0×10^{-15} metre apart. What is their mutual electric potential energy?

Soln.

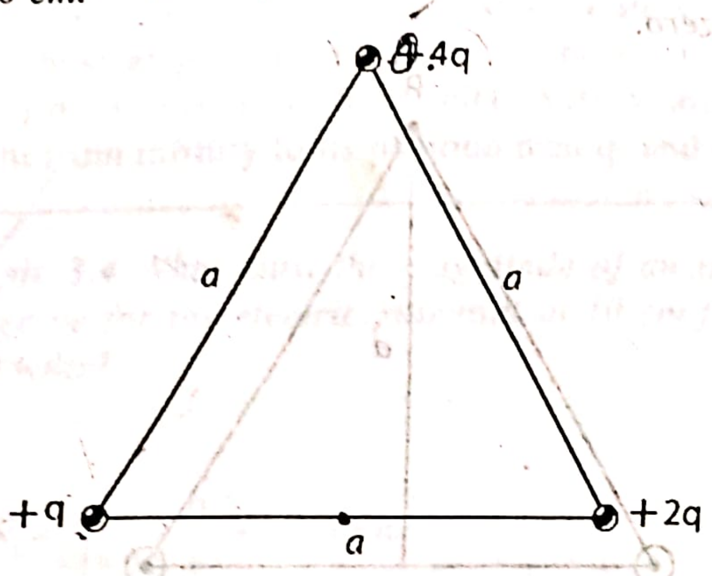
$$U=\frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r} = \frac{(9.0\times10^9\,nt-m^2/coul^2)(1.6\times10^{-19}coul)^2}{6.0\times10^{-15}m}$$

$$= 3.8 \times 10^{-14} \text{ Joule}$$

$$= \frac{3.8 \times 10^{-14}}{1.6 \times 10^{-19}} \text{ eV}$$

$$= 2.4 \times 10^{5} \text{ eV.}$$

Example 3.8 Three charges are arranged as in the figure. What is their mutual potential energy? Assume that $q = 1.0 \times 10^{-7}$ coul and $a = 10$ cm.



Soln.

The total energy of the configuration is the sum of the energies of each pair of particles. Therefore

$$U = U_{12} + U_{13} + U_{23}$$

$$= \frac{1}{4\pi\epsilon_0}\left[\frac{(-4q)(+q)}{a} + \frac{(+q)(+2q)}{a} + \frac{(+2q)(-4q)}{a}\right]$$

$$= \frac{1}{4\pi\epsilon_0}\left[\frac{-10q^2}{a}\right]$$

$$= -\frac{(9.0 \times 10^{9} \text{ nt} - \text{m}^2/\text{coul}^2)(10)(1.0 \times 10^{-7} \text{coul})^2}{0.10 \text{ m}}$$

$$= -9.0 \times 10^{-3} \text{ Joule.}$$

***Example 3.9** Two metal spheres are 3.0 cm in radius and carry charges of $+1.0 \times 10^{-8}$ coul and -3.0×10^{-8} coul respectively, assumed to be uniformly distributed. If their centre are 2.0m apart, calculate (a) the potential of the point halfway between their centres and (b) the potential of each sphere.*

Soln.

The charges may be assumed to be located at the centres of each sphere. The midway point will be 1.0 metre from either centres. The potential at the mid-point is then given by

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r_1} + \frac{1}{4\pi\epsilon_0}\frac{q_2}{r_2}$$

But $r_1 = r_2 = 1.0$ metre.

$$\therefore \quad V = \frac{1}{4\pi\epsilon_0}\left[\frac{q_1 + q_2}{r}\right]$$

$$= \frac{(9.0\times10^9\,\text{N}-\text{m}^2/\text{c}^2)(1\times10^{-8} - 3\times10^{-8})}{1.0\text{m}}$$

$$= \frac{-18\times10^9\times10^{-8}}{1.0} = -180 \text{ volts.}$$

(b) The potential at the surface of sphere 1 is due to its own charge q_1 plus that due to charge q_2 of sphere 2 at a distance r.

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{R} + \frac{1}{4\pi\epsilon_0}\frac{q_2}{r} \qquad R = 3 \text{ cm} = 0.03 \text{ m}, \quad r = 2\text{m}$$

$$= (9 \times 10^9 \text{ N-m}^2/\text{coul}^2)\left[\frac{1\times10^{-8}\text{c}}{0.03} + \frac{-3\times10^{-8}\text{c}}{2.0\text{m}}\right]$$

$= 2864$ volts.

The potential at the surface of sphere 2 is due to its own charge q_2 plus that due to charge q_1 of sphere 1 at a distance r.

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_2}{r} + \frac{1}{4\pi\epsilon_0}\frac{q_1}{R}$$

$$= (9 \times 10^9 \text{ N-m}^2/\text{coul}^2) \left[\frac{-3\times10^{-8}\text{c}}{0.03\text{m}} + \frac{1\times10^{-8}\text{c}}{2\text{m}} \right]$$

$$= -8955 \text{ volts.}$$

Example 3.10 Calculate (i) the electric potential established by the nucleus of a hydrogen atom at the mean distance of the circulating electron ($r = 5.3 \times 10^{-11}$m), (ii) the electric potential energy of the atom when the electron is at its radius and (iii) the kinetic energy of the electron assuming it to be moving in a circular orbit of this radius centred on the nucleus. (iv) how much energy is required to ionize the hydrogen atom? Express all energies in electron-volts.

Soln.

(i) The electric potential is given by

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r} = \frac{(9\times10^9 \text{ N}-\text{m}^2/\text{c}^2)(1.6\times10^{-19}\text{c})}{5.3\times10^{-11}\text{m}}$$

$$= 27.1 \text{ volts.}$$

(ii) Electric potential energy of the atom is given by

$$U = qV = -e.V = -(1.6 \times 10^{-19}\text{c})(27.1 \text{ volts})$$

$$= -\frac{(1.6\times10^{-19}\text{c})(27.1\text{volts})}{1.6\times10^{-19}} \text{eV}$$

$$= -27.1 \text{ eV.}$$

(iii)From the relation of electrostatic force balancing centripetal force, we obtain

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0}\frac{e^2}{r^2}$$

$$\text{or, } mv^2 = \frac{e^2}{4\pi\epsilon_0 r^2} = 27.1 \text{ eV}$$

$$\therefore \text{ K. E.} = \frac{1}{2}mv^2 = 13.55 \text{ eV.}$$

(iv) Total energy = kinetic energy + potential energy

$= 13.55 \text{ eV} - 27.1 \text{ eV}$

$= - 13.55 \text{ eV}.$

This is the energy required to ionize the hydrogen atom.

***Example 3.11** A spherical drop of water carrying a charge of 3×10^{-6} coul has a potential of 500 volts at its surface (i) What is the radius of the drop? (ii) If two such drops of the same charge and radius combine to form a single spherical drop, what is the potential at the surface of the new drop formed?*

Soln.

(i) From $V = \dfrac{1}{4\pi\epsilon_0}\dfrac{q}{r}$, we have

$$r = \frac{1}{4\pi\epsilon_0}\frac{q}{V} = (9 \times 10^9 \text{ N-m}^2/c^2) \times \frac{3\times10^{-6}\text{C}}{500\text{V}}$$

$= 54$ m.

(ii) Assuming the drops to be incompressible, the volume of the new drop will be twice the volume of the small drop.

or, $\dfrac{4}{3}\pi R^3 = 2 \times \dfrac{4}{3}\pi r^3$

where R is the radius of the new drop.

$\therefore \; R = 2^{1/3} . r$

the charge on the new drop, Q = 2q.

$\therefore$ potential at the surface of the new drop

$$= \frac{1}{4\pi\epsilon_0}\frac{Q}{R} = \frac{1}{4\pi\epsilon_0}\frac{2q}{2^{1/3}.r} = \frac{1}{4\pi\epsilon_0}\frac{2^{2/3}q}{r}$$

$= 500 \times 2^{2/3}$ volts

$= 790$ volts.

3.10 Equipotential surfaces

If a surface can be imagined in such a way that it is everywhere at right angles to an electric field, then any path connecting two points on that surface is always at right angles to the field. The potential difference between the two points is given by

$$V_B - V_A = \int_A^B E.dl = \int_A^B E \cos 90^o \, dl = 0$$

In other words, all points on that surface are at the same potential. Such a surface is called an *equipotential surface.* Alternately, the locus of points, all of which have the same electric potential, is called an equipotential surface.

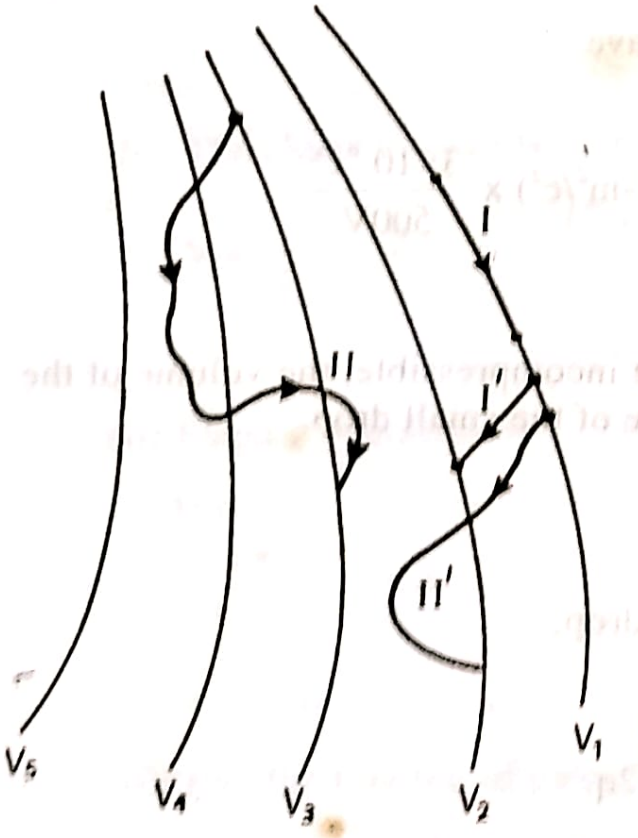


Fig. 3.9

A family of equipotential surfaces, each surface corresponding to a different value of the potential is shown in Fig. 3.9. The work to move a charge along path I and II is zero because all these paths begin and end on the same potential. But, on the other hand, the work to move a charge along paths I' and II' is not zero. Moreover, the amount of work done is also same because the initial and final potentials are identical; paths I' and II' connect the same pair of equipotential surfaces.

For an isolated point charge the potential at any point is given by

$$V = \frac{1}{4\pi\epsilon_0} \frac{q}{r}$$

Thus for an isolated point charge the equipotential surfaces are spheres concentric with the point charge (Fig. 3.10). It can be seen

that the electric field strength **E** is everywhere normal to the equipotential surface.

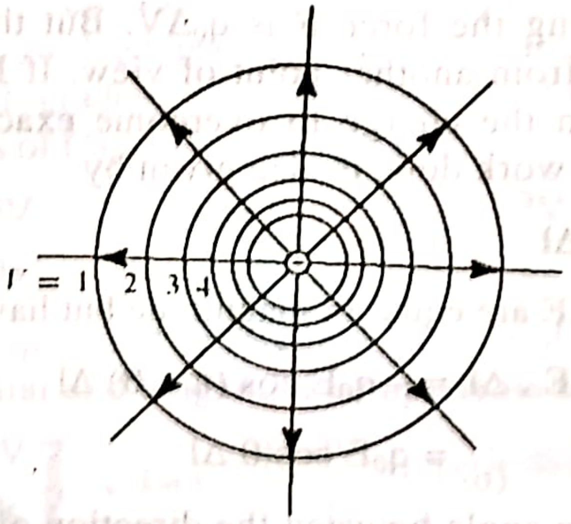


Fig. 3.10

Calculation of the electric field (E) from the potential (V)

It was mentioned at the very begining of this chapter that the electric field around a charged particle can be described by both electric field strength (E) and the electric potential (V). In Art. 3.3 we have shown how to calculate V from E. In this section, we propose to go the other way, that is, to find the electric field when we know the potential.

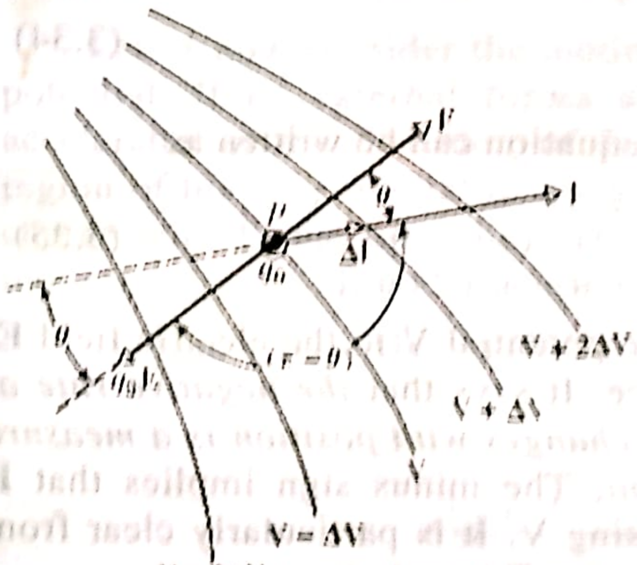


Fig. 3.11

Fig. 3.11 shows the intersection with the plane of the figure of a family of equipotential surfaces, differing in potential by the amount dV. The figure also shows the electric field **E** at a point P on the equipotential surface V. From the definition of equipotential surface, **E** should be at right angles to the equipotential surface through P.

Let us move a test charge q_0 from P along the path marked $\Delta \mathbf{l}$ to the equipotential surface marked $V+\Delta V$. The work that must be done by the agent exerting the force **F** is $q_0 \Delta V$. But the work done can also be calculated from another point of view. If **F** is the force that must be exerted on the charge to overcome exactly the electriclal force $q_0\mathbf{E}$, then the work done is also given by

$$\Delta W = \mathbf{F} \cdot \Delta \mathbf{l}$$

Since **F** and $q_0\mathbf{E}$ are equal in magnitude but have opposite signs,

$$\Delta W = - q_0\mathbf{E} \cdot \Delta \mathbf{l} = - q_0 E \cos(\pi - \theta)\, \Delta l$$

$$= q_0 E \cos\theta\, \Delta l$$

where $(\pi - \theta)$ is the angle between the direction of the force (**E**) and the direction of the displacement ($\Delta \mathbf{l}$).

Since the two expressions for work done must be equal,

$$q_0 \Delta V = q_0 E \cos\theta\, \Delta l$$

$$\text{or,} \quad E \cos\theta = \frac{\Delta V}{\Delta l} \tag{3.33}$$

Now $E \cos\theta$ is the component of **E** in the direction of $-l$ in Fig. 3.11. Therefore, the quantity $-E\cos\theta$, which we call E_l, would be the component of **E** in the $+l$ direction. We then obtain

$$E_l = -\frac{\Delta V}{\Delta l} \tag{3.34}$$

In the differential limit this equation can be written as

$$E_l = -\frac{dV}{dl} \tag{3.35}$$

Eqn. 3.35 relates the electric potential V to the electric field **E**, and is of fundamental importance. It says that *the negative rate at which the potential at any point changes with position is a measure of the electric field at that point.* The minus sign implies that **E** points in the direction of decreasing V. It is particularly clear from eqn. 3.35, that the appropriate unit for **E** is volts/metre.

The value of dV/dl at a point in a non-uniform field is called the *potential gradient* at the point in the direction of increasing distance. In other words, the resolved part of **E** in any direction is equal to the negative potential gradient in that direction. In particular, if V is expressed in terms of Cartesian coordinates x, y, z, then

$$E_x = -\frac{\partial V}{\partial x}; \quad E_y = -\frac{\partial V}{\partial y}; \quad E_z = -\frac{\partial V}{\partial z} \tag{3.35a}$$

If V can be specified in terms of plane polar coordinates (r, θ) only, then the radial and tangential components of **E** are

$$E_r = -\frac{\partial V}{\partial r}; \quad E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} \tag{3.35b}$$

3.12 The motion of charged particles in electric field

Let us consider a particle of mass m and charge q moving with velocity **v** in an electric field **E**. We shall assume that the charge is small enough not to affect the field in which it is placed or that the field can be contained constant by its sources. If no other forces act on the particle, then its equation of motion is

$$q\mathbf{E} = m\mathbf{a}; \quad \text{or,} \quad \mathbf{a} = \frac{q\mathbf{E}}{m}$$

where **a** is the acceleration of the particle.

Let us now consider the motion between two points in terms of potential. If no external forces act, a positively charged particle accelerates in the direction of **E** and in doing so moves from a region of higher potential to a region of lower potential (a negative charge will do the opposite). Thus the loss in potential energy in moving from A to B under action of **E** alone (Fig. 3.12) is

$$U_{AB} = U_A - U_B = q(V_A - V_B)$$

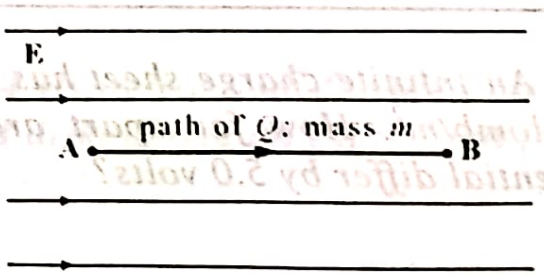


Fig. 3.12

This loss in potential energy is compensated by a gain in kinetic energy of the particle. For a charge moving along a line of **E**, the potential difference $V_A - V_B$ is simply the line integral of –E from B to A. Hence

$$U_A - U_B = q\,(V_A - V_B) = q\int_A^B + E\,dr$$

$$= \int_A^B madr$$

$$= \int_A^B mvdv \qquad \text{writing } a = v\frac{dv}{dr}$$

$$= \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 \tag{3.36}$$

The total mechanical energy $(U + \frac{1}{2}mv^2)$ remains unchanged at A and B and is thereby conserved. Although eqn. 3.36 was derived for a charge moving along a line of force, the velocity perpendicular to **E** will be unchanged and hence the relation applies to any trajectory.

In atomic and nuclear particles, charged particles are accelerated to high kinetic energies by the application of electric fields. Rather than quote the energies of such particles in joules, it is more convenient to express it in units of *electron-volt. One electron-volt, abbreviated eV, is the kinetic energy gained or lost by an electron when it is accelerated or retarded by a potential difference of one volt.* For higher energies 1 MeV (= 10^6 eV) and 1 GeV (= 10^9 eV) are often used as units although none of these is an SI unit. The latter remains the joule, related to eV by 1eV = 1.6×10^{-19} J.

Example 3.12 *An infinite charge sheet has a surface density σ of 1.0×10^{-7} coulomb/m². How far apart are the equipotential surfaces whose potential differ by 5.0 volts?*

Soln.

Electric field intensity is given by

$$E = \frac{\sigma}{2\epsilon_0} = \frac{1.0\times10^{-7}\,\text{coulomb/m}^2}{2\times8.9\times10^{-12}\,\text{coulomb}^2/\text{N}-\text{m}^2} = 5.610^3\,\text{N/coulomb}$$

Let the equipotential be dS metres apart. Then

$$E.dS = dV$$

or, $$dS = \frac{dV}{E} = \frac{5.9\ \text{volt}}{5.6\times10^3\,\text{N/coulomb}}$$

$$= 0.89 \times 10^{-4}\ \text{metres}$$

$$= 0.89\ \text{mm.}$$

***Example 3.13** Using eqn. 3.35, calculate E(r) for a point charge q. Assume that V(r) is given by*

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}$$

Soln.

From symmetry, E must be directed radially outward for a positive point charge. Consider a point P in the field a distance r from the charge. It is clear that dV/dl at P will have its greatest value if l points in the same direction as r. Thus, from eqn. ......

$$E = -\frac{dV}{dr} = -\frac{d}{dr}\left[\frac{1}{4\pi\epsilon_0}\frac{q}{r}\right]$$

$$= -\frac{q}{4\pi\epsilon_0}\frac{d}{dr}\left(\frac{1}{r}\right) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$$

The result agrees exactly with eqn. 1.7, as it must.

***Example 3.14** Using eqn. 3.8 for the potential on the axis of a uniformly charged disk, derive an expression for the electric field at axial points.*

Soln.

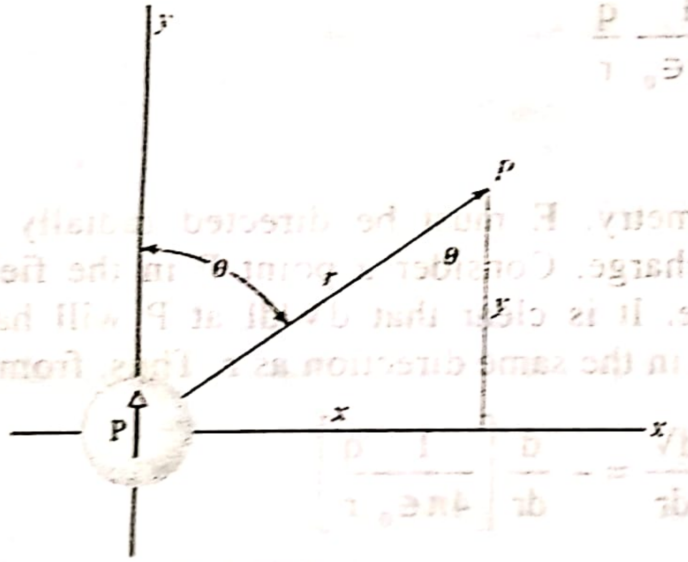
From symmetry, E must lie along the axis of the disk (the r axis). Using eqn. 3.8 we have

$$E_r = -\frac{\partial V}{\partial r} = -\frac{\sigma}{2\epsilon_0}\frac{d}{dr}\left[\left(a^2 + r^2\right)^{1/2} - r\right]$$

$$= \frac{\sigma}{2\epsilon_0}\left(1 - \frac{r}{\sqrt{a^2 + r^2}}\right)$$

The same expression can be obtained by direct integration using coulomb's law.

***Example 3.15** Fig. below shows a (distant) point P in the field of a dipole located at the origin of an xy-coordinate system. Calculate E as a function of position.*



Soln.

From symmetry, E, for points in the plane of Fig. 3.14, lies in this plane. Thus it can be expressed in terms of its components E_x and E_y. We shall first express the potential function in rectangular coordinates rather than polar coordinates, making use of

$$r = (x^2 + y^2)^{1/2} \text{ and } \cos\theta = \frac{y}{\left(x^2 + y^2\right)^{1/2}}$$

The resultant expression for V is given by eqn. 3.17.

$$V = \frac{1}{4\pi\epsilon_0}\frac{p\cos\theta}{r^2}$$

$$= \frac{p}{4\pi\epsilon_0}\frac{y}{(x^2+y^2)^{3/2}}$$

We find E_y from eqn. 3.35a, recalling that x is to be treated as a constant in this calculation.

$$E_y = -\frac{\partial V}{\partial y} = -\frac{p}{4\pi\epsilon_0}\frac{(x^2+y^2)^{3/2} - y\frac{3}{2}(x^2+y^2)^{1/2}(2y)}{(x^2+y^2)^3}$$

$$= -\frac{p}{4\pi\epsilon_0}\frac{x^2-2y^2}{(x^2+y^2)^{5/2}}$$

Putting x = 0 describes points along the dipole axis (that is, the y axis), and the expression for E_y reduces to

$$E_y = \frac{1}{4\pi\epsilon_0}\frac{2p}{y^3}$$

Putting y = 0 in the expression for E_y describes points in the median plane of the dipole and the expression for E_y reduces to

$$E_y = -\frac{1}{4\pi\epsilon_0}\frac{p}{x^3}$$

The component E_x can also be found from eqn. ......, recalling that y is to be taken as a constant during this calculation:

$$E_x = -\frac{\partial V}{\partial x} = -\frac{py}{4\pi\epsilon_0}\left(-\frac{3}{2}\right)(x^2+y^2)^{-\frac{5}{2}}(2x)$$

$$= \frac{3p}{4\pi\epsilon_0}\frac{xy}{(x^2+y^2)^{5/2}}$$

As expected, E_x vanishes both on the dipole axis (x = 0) and in the median plane (y = 0).

***Example 3.16** The potential in the region of space near the point (– 2, 4, 6m) is V = $80x^2+60y^2$ V. What are the three components of the electric field at that point?*

Soln.

We have

$$E_x = -\frac{\partial V}{\partial x} = -160x = 320 \text{ V/m.}$$

$$E_y = -\frac{\partial V}{\partial y} = -120y = -480 \text{ V/m.}$$

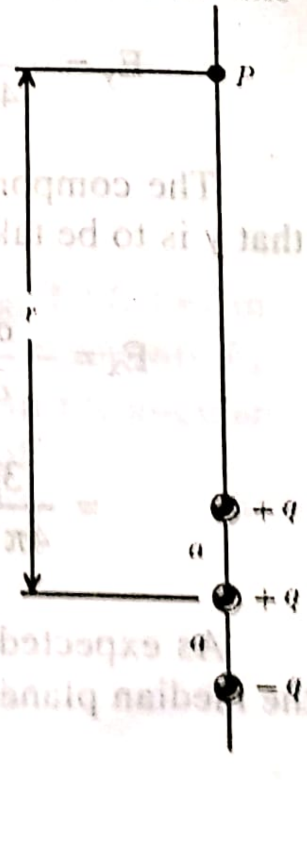
$$E_z = -\frac{\partial V}{\partial z} = 0.$$

***Example 3.17** Show that the potential at a point on the axis of charges as shown in the figure is given by*

$$V(r) = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r} + \frac{2qa}{r^2}\right) \quad \text{for } r >> a$$

Soln.

It is clear from the figure that the charges +q and –q, separated by a distance 2a, formed an electric dipole. The potential due to this dipole at a point P along the axis of the dipole is



$$V_1 = \frac{1}{4\pi\epsilon_0}\frac{p}{r^2} = \frac{1}{4\pi\epsilon_0}\frac{q.2a}{r^2} \qquad [p = q.2a]$$

The potential at P due to the charge +q,

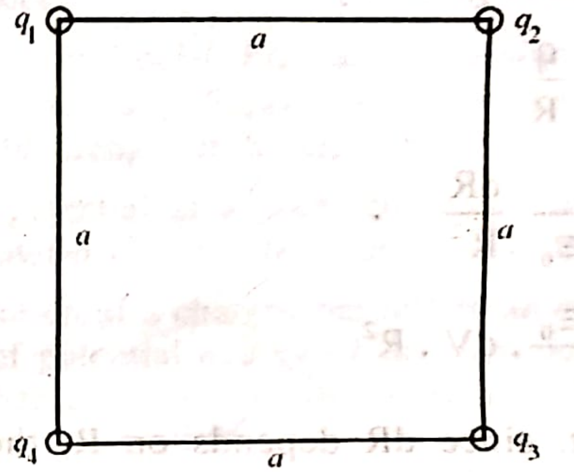
$$V_2 = \frac{1}{4\pi\epsilon_0}.\frac{q}{r}$$

∴ the resultant potential,

$$V = V_1 + V_2 = \frac{1}{4\pi\epsilon_0}\left(\frac{2aq}{r^2} + \frac{q}{r}\right)$$

***Example 3.18** What is the electric potentia... ne charge configuration of the figure below. Given, q_1 = ... v of the $q_2 = -2.0 \times 10^{-8}C$, $q_3 = +3.0 \times 10^{-8}C$, $q_4 = +2.0 \times 10^{-8}C$ an... $10^{-8}C$, ... m.*

Soln.



$$U = U_{12} + U_{13} + U_{14} + U_{23} + U_{24} + U_{34}$$

$$= \frac{1}{4\pi\epsilon_0}\left[\frac{q_1q_2}{a} + \frac{q_1q_3}{\sqrt{2}.a} + \frac{q_1q_4}{a} + \frac{q_2q_3}{a} + \frac{q_2q_4}{\sqrt{2}.a} + \frac{q_3q_4}{a}\right]$$

$$= 9 \times 10^9\left[\frac{(1.0)(-2.0)}{1} + \frac{(1.0)(3.0)}{\sqrt{2}.1} + \frac{(1.0)(2.0)}{1}\right.$$

$$\left. + \frac{(-2.0)(3.0)}{1} + \frac{(-2.0)(2.0)}{\sqrt{2}.1} + \frac{(3.0)(2.0)}{1}\right] \times 10^{-16}\ \text{J.}$$

$$= -6.4 \times 10^{-7}\text{J.}$$

***Example 3.19** Consider a point charge with $q = 1.5 \times 10^{-8}$ coulomb. (i) What is the radius of the equipotnetial surface having a potential of 30 volts. (ii) Are surfaces whose potentials differ by a constant amount (say 1.0 volt) evenly spaced in radius?*

Soln.

(i) Radius can be calculated from the relation

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{R}$$

$$\text{or, } R = \frac{1}{4\pi\epsilon_0}\frac{q}{V}$$

$$= \frac{(9\times10^9 \text{ N} - \text{m}^2/\text{coulomb}^2)(1.5\times10^{-8}\text{ coul})}{(30 \text{ volts})}$$

$= 4.5$ metres.

The potential is given by

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{R}$$

$$\therefore\ dV = -\frac{q}{4\pi\epsilon_0}\frac{dR}{R^2}$$

$$\text{or, } dR = -\frac{4\pi\epsilon_0}{q}\,.\,dV\,.\,R^2$$

It is clear that, since dR depends on R, the surfaces whose potential differ by a constant amount are not evenly spaced.

EXERCISES

1. Define potential at a point in an electric field. Find an expression for the potential at a point in an electric field due to a point charge.

2. Define Electric field strength and electric potential. Show that the potential at a point r is given by $V = -\int_{\infty}^{r} \mathbf{E}\,.\,d\mathbf{l}$ where the symbols have their usual meanings.

3. Do electrons tend to go to regions of high potential or of low potential?

4. Dose the amount of work per unit charge required to transfer electric charge from one point to another in an electrostatic field depend on the amount of charge transferred?

5. Define electric potential and electric potential energy.

6. Find the electric potential at a point on the axis at a distance r from the centre of a uniformly charged circular disk whose surface charge density is σ.

7. Determine the potential at a distance r from the centre of a uniformly charged conducting sphere of radius R for (i) $r > R$, (ii) $r = R$ and (iii) $r < R$. The total charge on the sphere is Q.

8. Define potential gradient and explain how you can calculate **E** if V is known through a certain region.

9. Show that the intensity at a point in an electric field is equal to the negative gradient of potential at the same point.

10. What is an equipotential surface? Can two different equipotential surfaces intersect?

11. What is an electric dipole? Find an expression for the potential at a point due to a dipole and discuss the different cases. Hence determine the electric field strength at the same point.

12. Calculate the potential at a point due to a charged ring and hence calculate the intensity at the same point due to this charged ring.

13. Discuss the motion of a charged particle in an electric field and show that the loss of potential energy of the particle in moving from one point to another point in an electric field is equal to the gain in kinetic energy of the particle.

14. Define electron-volt. How would a proton-volt compare with an electron-volt? The mass of a proton is 1840 times that of an electron.

15. What is the electric potential at the surface of a gold nucleus? The radius is 6.6×10^{-15} metre and the atomic number Z = 79. **[1.7×10^7 volts]**

16. What is the potential at the point midway between the charges +2μC and +5μC which are 6m apart? **[21 kV]**

17. Three equal charges of +6nC are located at the corners of an equilateral triangle whose sides are 12 cm long. Find the potential at the centre of the base of the triangle. **[2320 V]**

18. A metal sphere 30 cm in radius is positively charged with 2 μC. Find the potential at the centre of the sphere, on the sphere, and at 1m from the centre of the sphere. **[60 kV on the surface or at the centre, 18kV at 1 metre from the centre]**

19. Two protons in a nucleus of U^{238} are 6.0×10^{-15} metre apart. What is their mutual electric potential energy? **[3.8×10^{-14}J or 0.24 MeV]**

20. Three charges of +q, –4q and +2q are placed at the corners of an equilateral triangle of side *a*. What is their mutual potential energy? Assume q = 1.0×10^{-7}C and a = 10 cm. **– 9.0×10^{-3}J]**

21. A charge of 10^{-8} coulomb can be produced by simple rubbing. To what potential would such a charge raise an insulated conducting sphere of 10 cm radius? **[900 volts]**

CHAPTER IV

CAPACITANCE

4.1 Capacitor

Two conductors of arbitrary shape, completely isolated from each other and their surroundings, form a capacitor. No matter what their shape, these conductors are called *plates.* When the capacitor is charged by connecting the plates to the opposite terminals of a battery, equal and opposite charges (say +q and −q) appear on the

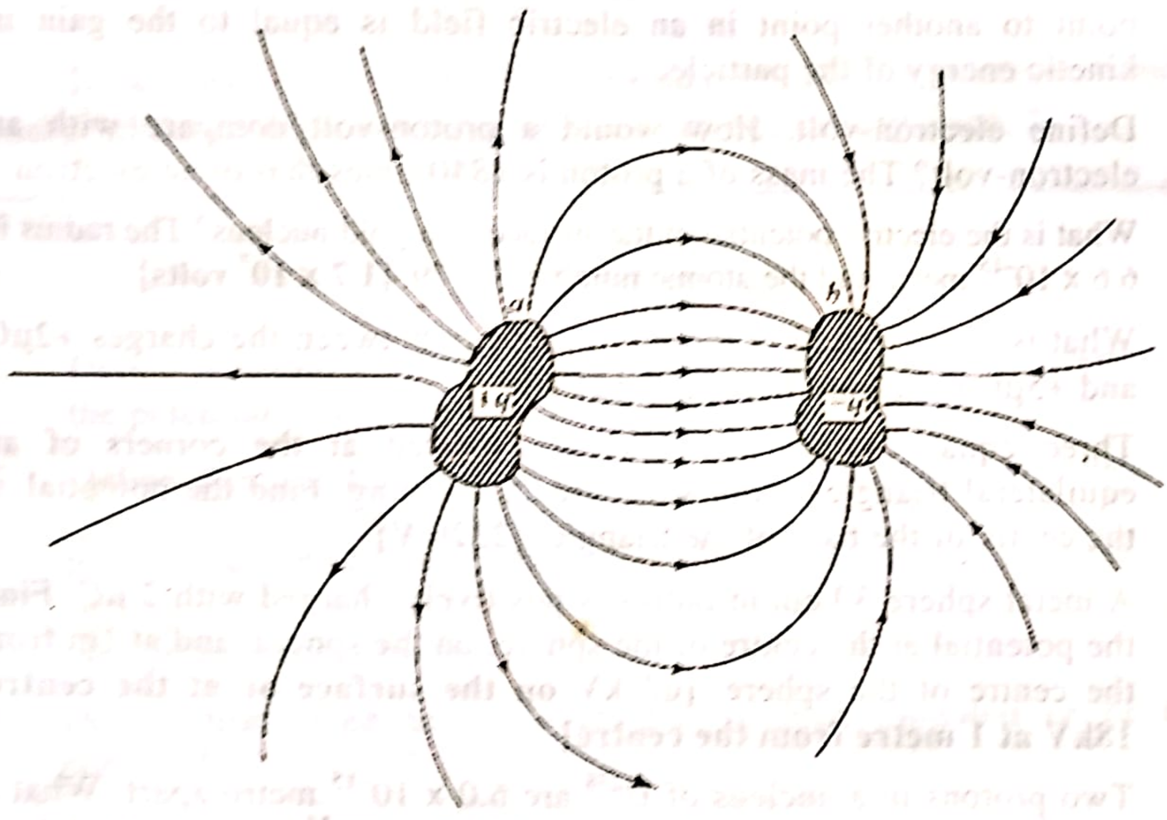


Fig. 4.1

two plates of the capacitor. By charge of a capacitor we mean the absolute value of the charge on either plate, the net charge on the capacitor being zero. The potential difference between the plates of the capacitor is the potential difference of the battery. Fig. 4.1 shows the general arrangement of a capacitor.

4.2 Capacitance ✓

The charge q of a capacitor is found to be directly proportional to the potential difference between the plates. Or,

$$q \propto V$$

$$\text{or, } q = CV \qquad \text{or, } C = \frac{q}{V} \qquad (4.1)$$

The proportionality constant C is called the *capacitance* of the capacitor. Its value depends on

(i) the geometry of each plate

(ii) the spatial relationship between the plates

and (iii) the medium in which the plates are immersed.

As can be seen from eqn. (4.1), the SI unit of capacitance is *coulomb per volt.* This unit occurs so often that it is given a special name – the *farad.*

1 farad (1F) = 1 coulomb per volt (1 C/V). So *if a potential difference of 1 volt is needed to give a capacitor a charge of 1 coulomb, then the capacitance of the capacitor is said to be 1 farad.*

The farad is a large unit. Submultiples of the farad, such as *microfarad* ($1\mu F = 10^{-6}F$) and the *picofarad* ($1pF = 10^{-12}F$) are more convenient units in practice.

Calculation of capacitance

Once the geometry of a capacitor is known, its capacitance can be calculated. Since different capacitors have different plate geometries, it is wise to develop a general plan to simplify the process involved in the calculation. In brief, the plan is

(i) assume a charge q on the plates

(ii) applying Gauss' law to calculate the electric field **E** between the plates in terms of the charge on the plates.

(iii) knowing **E**, calculate the potential difference V between the plates.

(iv) calculate C from $C = \frac{q}{V}$

The calculation of the electric field and the potential difference may be simplified by making certain assumptions. These are,

(a) calculating the electric field:

The electric field is related to the charge on the plates by Gauss' law:

$$\epsilon_0 \int \mathbf{E} . d\mathbf{A} = q \quad (4.2)$$

Here q is the charge enclosed by the Gaussian surface, and the integral is carried out over that surface. Only those cases will be considered in which the Gaussian surface are such that whenever electric flux passes through it, the electric field **E** and **dA** will point in the same direction. Eqn. 4.2 then reduces to

$$q = \epsilon_0 E A$$

in which A is the area of that part of the Gaussian surface through which the flux passes. For convenience the Gaussian surface is so drawn that it completely encloses the charge on the positive plate.

(b) Calculating the potential difference:

The potential difference between the plates is related to the electric field **E** by the relation

$$V_f - V_i = -\int_i^f \mathbf{E} . d\mathbf{S} \quad (4.3)$$

the integral being evaluated along any path that starts on one plate and ends on the other. One should always choose a path that follows an electric field line from the positive plate to the negative plate as shown in Fig. 4.2, since the vectors **E** and **dS** point in the same direction along this path. It, therefore, follows that the quantity $V_f - V_i$ is negative. Since we are looking for V, the *absolute value* of the potential difference between the plates, we can set $V_f - V_i = -V$. Eqn. 4.3 then becomes

$$V = \int_{+}^{-} E\,dS$$

in which the + and the – signs remind us that our path of integration starts on the positive plate and ends on the negative plate.

The electric field E between the plates is the sum of the fields due to the two plates *i.e.*, $E = E_+ + E_-$ where E_+ is the field due to charges on the positive plate while E_- is that due to charges on the negative plate. By Gauss' law both E_+ and E_- are proportional to q so that E is also proportional to q. By eqn. 4.3. V is also proportional to q. This means that if q is doubled, E and V are also doubled. Because V is proportional to q, the ratio q/V is a constant and is independent of q.

(i) *Capacitance of a parallel-plate capacitor*

A parallel-plate capacitor formed of two parallel conducting plates of area A and separated by a distance d is shown in Fig. 4.2. If the plates are connected to the opposite terminals of a battery, then a charge +q appears on one plate and a charge –q on the other. If d is small enough compared to the plate dimensions, the electric field strength E between the plates will be uniform, which means that the lines of force will be parallel and evenly spaced. According to the laws of electro-magnetism, there should be some *fringing* or curving of the lines at the edges of the plates; for small enough d it can be neglected for the present purpose.

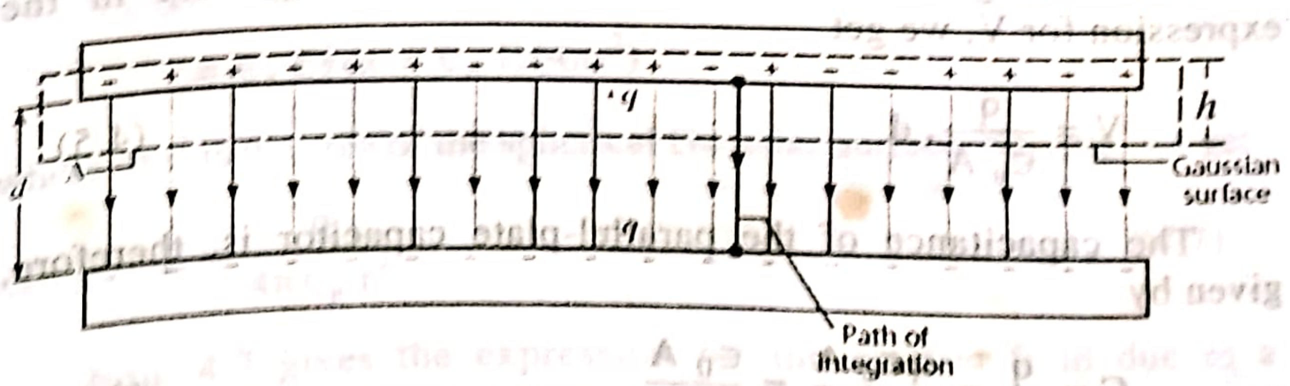


Fig. 4.2

Let us imagine a Gaussian surface of height h closed by plane caps of area A of the same shape and size of the capacitor plates. Because the electric field inside a conductor carrying a static charge is zero, the flux of **E** for the part of the Gaussian surface that lies inside the top capacitor plate is also zero. The flux **E** through the wall of the Gaussian surface is zero because, to the extent that the fringing of the lines of force can be neglected, **E** lies in the wall. Thus the only part of the Gaussian surface which contributes to the electric flux is the Gaussian surface that lies between the plates. Here **E** is constant and according to Gauss' law

$$\phi_E = \oint \mathbf{E}.d\,\mathbf{s} = E.\ A = \frac{q}{\epsilon_0}$$

$$\text{or,} \qquad \epsilon_0\, EA = q$$

$$\text{or,} \qquad E = \frac{q}{\epsilon_0 A} \tag{4.4}$$

The potential difference V between the plates can be obtained from eqn. 4.3. Or

$$V = \int_{+}^{-} \mathbf{E}.d\,\mathbf{l} = \int_{+}^{-} Edl = E\int_{+}^{-} dl = Ed.$$

since **E** is constant and can be taken outside the integral and $\int dl$ is simply the plate separation d.

Substituting the value of E as given by eqn. 4.4 in the expression for V, we get

$$V = \frac{q}{\epsilon_0 A}.\, d \tag{4.5}$$

The capacitance of the parallel-plate capacitor is, therefore, given by

$$C = \frac{q}{V} = \frac{q\epsilon_0 A}{qd} = \frac{\epsilon_0 A}{d} \tag{4.6}$$

As can be seen from eqn. 4.6, the capacitance does indeed depend only on geometrical factors, namely, the plate area A and the plate separation d.

(ii) Capacitance of a spherical capacitor

Fig. 4.3 shows a central cross-section of a capacitor that consists of two concentric spherical shells of radii *a* and *b*. As a Gaussian surface let us draw a sphere of radius *r* concentric with the two shells. Applying Gauss' law to this surface we obtain

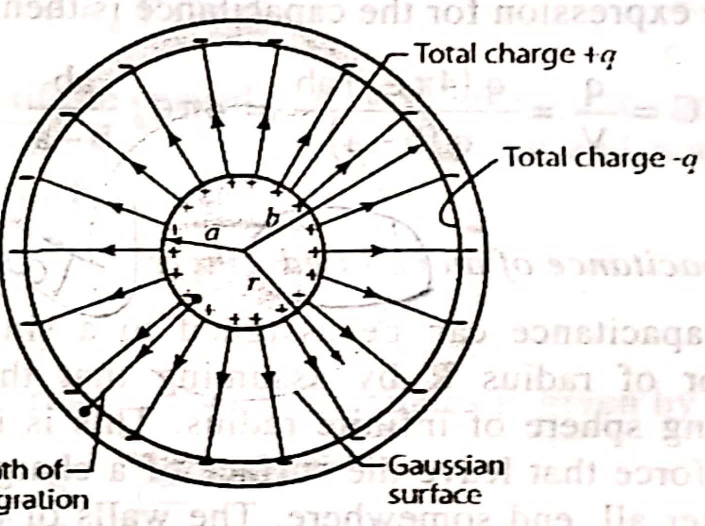


Fig. 4.3

$$q = \epsilon_0 \oint \mathbf{E} \cdot d\mathbf{s} = \epsilon_0 \oint E ds$$

$$= \epsilon_0 E \oint ds = \epsilon_0 E (4\pi r^2)$$

where $4\pi r^2$ is the area of the spherical Gaussian surface. Solving, we get

$$E = \frac{q}{4\pi\epsilon_0 r^2} \qquad (4.7)$$

Eqn. 4.7 gives the expression for the electric field due to a uniform spherical charge distribution.

The expression for the potential difference between the two concentric spheres is given by

$$V = \int_a^b E\,ds = \int_a^b \frac{q}{4\pi\epsilon_0 r^2}\cdot dr$$

$$= \frac{q}{4\pi\epsilon_0}\int_a^b \frac{dr}{r^2} = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{a}-\frac{1}{b}\right)$$

$$= \frac{q}{4\pi\epsilon_0}\cdot\frac{b-a}{ab} \qquad (4.8)$$

In deriving eqn. 4.8 we have used the fact that here $ds = dr$.

The expression for the capacitance is then given by

$$C = \frac{q}{V} = \frac{q\,(4\pi\epsilon_0)\,ab}{q\,(b-a)} = 4\pi\epsilon_0\frac{ab}{b-a} \qquad (4.9)$$

(iii) Capacitance of an isolated sphere

A capacitance can be assigned to a single isolated spherical conductor of radius R by assuming that the *missing plate* is a conducting sphere of infinite radius. This is necessary because the lines of force that leave the surface of a charged isolated conductor must, after all, end somewhere. The walls of the room in which the conductor is housed can effectively serve as the sphere of infinite radius.

Now the capacitance of a spherical capacitor as given by eqn. 4.9, is

$$C = 4\pi\epsilon_0\frac{ab}{b-a} = 4\pi\epsilon_0\frac{a}{1-a/b}$$

If the second sphere is of infinite radius, then $b \to \infty$. Substituting R for a, we obtain

$$C = 4\pi\epsilon_0 R \qquad (4.10)$$

(iv) Capacitance of a cylindrical capacitor

Fig. 4.3 also serves to show a cross-section of a cylindrical capacitor of length *l* formed by two co-axial cylinders of radii *a* and

b. The length of the capacitor is assumed to be much greater than its radius, *i.e.*, $l >> b$ so that fringing of the lines of force at the ends (edge effect) can be ignored for the purpose of calculating the capacitance. Each plate contains a charge of magnitude q.

As a Gaussian surface let us construct a coaxial cylinder of radius r and length l closed by end caps. Applying Gauss' law we then obtain

$$q = \epsilon_0 \oint \mathbf{E} . d\,\mathbf{s} = \epsilon_0 \oint Eds$$

$$= \epsilon_0 E \oint ds \quad = \epsilon_0 E (2\pi r l)$$

where $2\pi r l$ is the area of the curved part of the Gaussian surface, the flux being entirely through the cylindrical surface and not through the end caps. Solving for E we get

$$E = \frac{q}{2\pi\epsilon_0 rl}$$

The potential difference between the plates is given by

$$V = \int_{+}^{-} Eds = \frac{q}{2\pi\epsilon_0 l}\int_a^b \frac{dr}{r}$$

$$= \frac{q}{2\pi\epsilon_0 l} ln . \frac{b}{a}$$

From the relation $C = \frac{q}{V}$, we then have

$$C = 2\pi\epsilon_0 \frac{l}{ln\left(\frac{b}{a}\right)} \tag{4.12}$$

As can be seen from eqn. 4.12 that, like a parallel-plate capacitor; the capacitance of a cylindrical capacitor depends only on geometrical factors, in this case l, b and a.

The capacitances of various capacitors derived in this section is summarized below.

Type of capacitor	Capacitance	Equation
Paralle-plate	$\in_0 \frac{A}{d}$	4.6
Spherical	$4\pi\in_0 \frac{ab}{b-a}$	4.9
Isolated sphere	$4\pi\in_0 R$	4.10
Cylindrical	$2\pi\in_0 \frac{l}{\ln(b/a)}$	4.12

It can be seen that every expression involves the constant $\in_0$ multiplied by a quantity that has the dimension of length.

Example 4.1 A plane-parallel capacitor has circular plates of radius r = 10.0 cm, separated by a distance d = 1.00mm. How much charge is stored on each plate when their electric potential difference has the value V = 100V?

Soln.

q = CV where C is the capacitance of the capacitor and is given by

$$C = \frac{\in_0 A}{d}$$

Here, $A = \pi r^2 = \pi\,(0.1\text{m})^2$

$= 3.14 \times 10^{-2}\text{m}^2$

$d = 1.00\text{ mm} = 1 \times 10^{-3}\text{m}$

$$\therefore\ C = \frac{(8.85\times10^{-12}\text{C}^2/\text{N.m}^2)(3.14\times10^{-2}\text{m}^2)}{1.00\times10^{-3}\text{m}}$$

$= 2.8 \times 10^{-10}\text{ F} = 280 \times 10^{-12}\text{F}$

$= 280\text{pF}.$

$\therefore\ q = CV = (2.8 \times 10^{-10}\text{F})(100\text{ V})$

$= 2.8 \times 10^{-8}$ coulomb

$= 28 \times 10^{-9}\text{ C} = 28\text{ nC}.$

4.1 ~ 4.7, 4.10, 4.13

***Example 4.2** The area of each plate of an air-filled parallel-plate capacitor is 1.1 x 10^8 metre^2. What must be the separation between the plates, if the capacitance is to be 1.0 farad?*

Soln.

$$C = \frac{\epsilon_0 A}{d}$$

Here, $A = 1.1 \times 10^8 m^2$

$C = 1$ F.

$$\text{or, } d = \frac{\epsilon_0 A}{C}$$

$$= \frac{(8.85\times10^{-12} F/m)(1.1\times10^8 m^2)}{1.0 F}$$

$$= 9.735 \times 10^{-4} m$$

$$= 9.735 \times 10^{-1} mm = 0.9735 mm.$$

***Example 4.3** How much charge is stored in a capacitor consisting of two concentric spheres of radii 30 and 31 cm if the potential difference is 500V?*

Soln.

$q = CV$ where

Here, $a = 30$ cm $= 0.30$ m

$b = 31$ cm $= 0.31$ m

$$C = 4\pi\epsilon_0 \frac{ab}{(b-a)}$$

$$= \frac{(4)(3.14)(8.85\times10^{-12} F/m)(0.30m)(0.31m)}{(0.31-0.30)m}$$

$$= \frac{10.34\times10^{-12}}{0.01} F = 1034 \times 10^{-12} F$$

$$= 1.034 \times 10^{-9} F$$

$$= 1.034 \text{ nF}$$

$$q = CV = (1.03 \times 10^{-9} F)(500V) = 517 \text{ nC}.$$

***Example 4.4** What is the capacitance of the Earth, viewed as an isolated conducting sphere of radius 6370 km?*

Soln.

$$C = 4\pi\epsilon_0 R$$
$$= (4)(3.14)(8.85 \times 10^{-12}\text{F/m})(6370 \times 10^3\text{m})$$
$$= 7.08 \times 10^{-4}\text{F}$$
$$= 708 \times 10^{-6} = 708\ \mu\text{F}.$$

Example 4.5 The space between the conductors of a long coaxial cable, used to transmit TV signals, has an inner radius a = 0.15mm and an outer radius b = 2.1 mm. What is the capacitance per unit length of this cable?

Soln.

The capacitance of a coaxial cable is given by (eqn. 4.12)

$$C = 2\pi\epsilon_0 \frac{l}{\ln(b/a)}$$ where l is the length of the cable.

Hence capacitance per unit length is

$$\frac{C}{l} = \frac{2\pi\epsilon_0}{\ln(b/a)}$$

Here, $b = 2.1$ mm $= 2.1 \times 10^{-3}$m

$a = 0.15$ mm $= 0.15 \times 10^{-3}$m.

$$= \frac{(2)(3.14)(8.85\times10^{-12}\text{F/m})}{\ln\left(\dfrac{2.1\times10^{-3}\text{m}}{0.15\times10^{-3}\text{m}}\right)}$$

$$= \frac{55.578\times10^{-12}}{2.64}\text{F/m}$$

$$= 21 \times 10^{-12}\text{F/m} = 21\text{pF}.$$

4.3 Capacitors in series and parallel

In analyzing electric circuits, very often it is desirable to know the *equivalent capacitance* of two or more capacitors that are connected in a certain way. By equivalent capacitance is meant the capacitance of a single capacitor that can be substituted for the combination with no change in the operation of the rest of the circuit. With such a replacement, the circuit can be simplified, so

that we can solve for unknown quantities in the circuit more easily. In a circuit, capacitors can be combined either in series or parallel.

(a) Capacitors connected in series

Fig. 4.4 shows three capacitors connected *in series* to a battery B, maintaining a potential difference of V volts across the left and right terminals of the series combination. The potential differences produced across the individual capacitors are V_1, V_2, and V_3, the corresponding capacitances being C_1, C_2 and C_3 respectively. The connection is said to be series if $V_1 + V_2 + V_3 = V$.

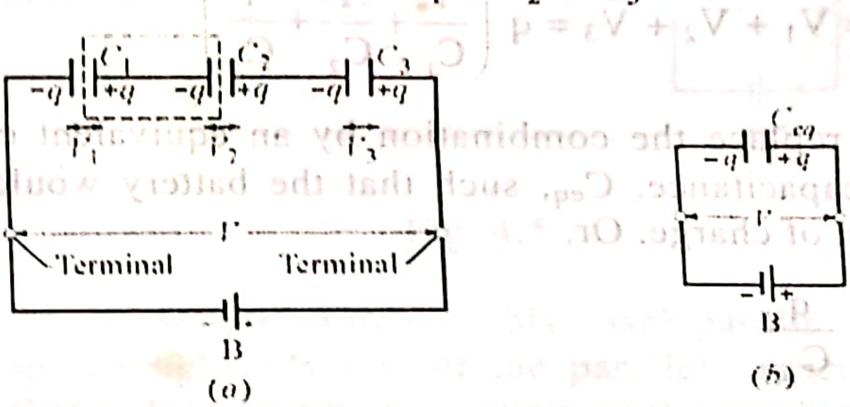


Fig. 4.4

Although the potential differences produced across the capacitors are different (depending on the individual capacitances), each capacitor has the same charge q, irrespective of its capacitance. To understand this let us consider the part of the circuit enclosed by the dashed line in the figure. Since this portion of the circuit is electrically isolated from the rest of the circuit, this may be regarded as a *floating* circuit. Let us assume that the battery puts a charge $-q$ on the left hand plate of C_1. Since a capacitor carries equal and opposite charges on its plates, a charge $+q$ appears on the right hand plate of C_1. Since the floating portion of the circuit is electrically isolated, initially it carries no net charge, and no charge can be transferred to it. Thus if a charge $+q$ appears on the right hand plate of C_1, then a charge $-q$ must appear on the left-hand plate of C_2. The presence of charge on the plates of the capacitors, not directly connected to the battery can be similarly explained. The result is that the left hand plate of *every* capacitor of the series connection carries a charge q of one sign while the right hand plate of every capacitor carries a charge of equal magnitude q but of opposite sign.

Another characteristic feature of series connection is that in going from one terminal to the other terminal, one must pass through *all* the circuit elements *in succession.*

For each individual circuit element we can write

$$V_1 = \frac{q}{C_1}, \quad V_2 = \frac{q}{C_2}, \quad \text{and} \quad V_3 = \frac{q}{C_3}$$

The potential difference of the series combination is then

$$V = V_1 + V_2 + V_3 = q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right) \qquad \text{(i)}$$

Let us replace the combination by an equivalent capacitor of equivalent capacitance. C_{eq}, such that the battery would move the same amount of charge. Or,

$$V = \frac{q}{C_{eq}} \qquad \text{(ii)}$$

Combing eqns. (i) and (ii) we obtain

$$C_{eq} = \frac{q}{V} = \frac{q}{q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right)}$$

$$\text{or,} \quad \frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \qquad \text{(4.13)}$$

Thus the equivalent capacitance is the reciprocal of the sum of the reciprocals of the individual capacitances.

(ii) Capacitors connected in parallel

Fig. 4.5 shows three capacitors connected in *parallel* to a battery B. The terminals of the battery are wired directly to the plates of the three capacitors. The connection is said to be parallel if the same potential difference V, which is the potential difference between the terminal of the battery, is applied across each of the capacitors. In going from one terminal to the other terminal, we can

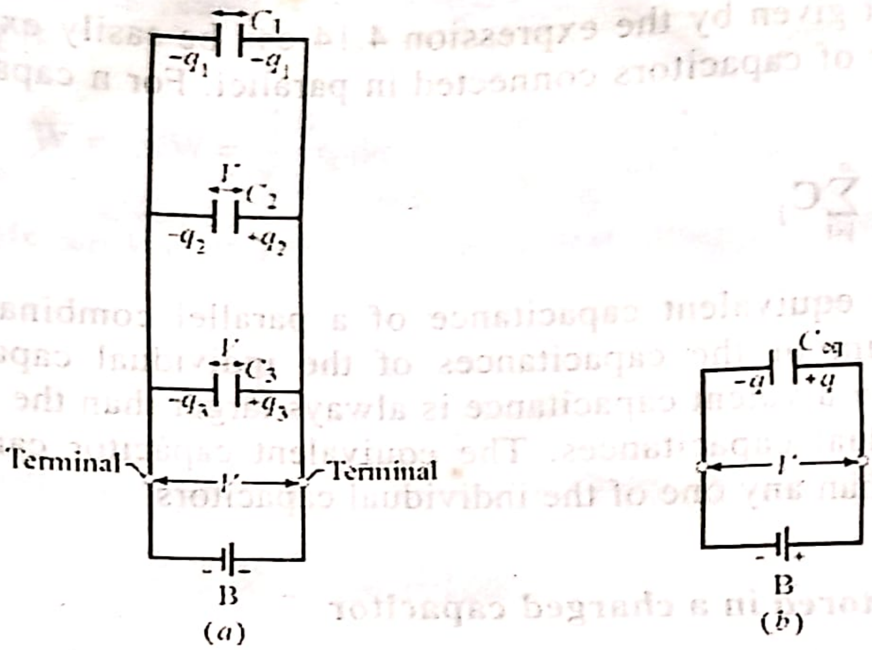


Fig. 4.5

take one of the several (three in this case) parallel paths, each of which goes through only one of the parallel capacitors. The total charge q that is delivered by the battery to the combination is shared among the capacitors. While calculating the equivalent capacitance C_{eq} of the single capacitor that can replace the combination we must remember that both the potential difference applied across the combination and the total charge q must not be changed.

For each capacitor we can write

$$q_1 = C_1V, \quad q_2 = C_2V, \quad \text{and} \quad q_3 = C_3V.$$

where q_1, q_2, and q_3 are the charges stored in the first, second and the third capacitors respectively and C_1, C_2 and C_3 are the corresponding capacitances of the capacitors.

The total charge on the parallel combination is then

$$q = q_1 + q_2 + q_3 = C_1V + C_2V + C_3V$$
$$= (C_1 + C_2 + C_3)\,V$$

The equivalent capacitance, with the same total charge q and the applied potential difference V as the combination is then

$$C_{eq} = \frac{q}{V} = \frac{(C_1 + C_2 + C_3)V}{V} = C_1 + C_2 + C_3 \quad (4.14)$$

The result given by the expression 4.14 can be easily extended to any number of capacitors connected in parallel. For n capacitors, we have

$$C_{eq} = \sum_{j=1}^{n} C_j$$

Thus the equivalent capacitance of a parallel combination is simply the sum of the capacitances of the individual capacitors. Note that the equivalent capacitance is always larger than the largest of the individual capacitances. The equivalent capacitor can store more charge than any one of the individual capacitors.

4.4 Energy stored in a charged capacitor

In chapter III, we have seen that any charge configuration has a certain electric potential energy U, equal to the work W that must be done by an external agent to assemble the charge configuration from its individual components, originally assumed to be infinitely far apart and at rest. Similarly, charging of a capacitor needs work to be done by an external agent. The process of charging an uncharged capacitor can be visualized as electrons being removed from one plate by some external agent. The electric field that builds up in the space between the plates will be in a direction that tends to oppose further transfer. Thus as charge accumulates on the capacitor plates, increasingly larger amounts of work will be required to transfer additional electrons. This work is stored in the form of electric potential energy U in the electric field between the plates. In practice this work is done by a battery at the expense of its store of chemical energy. The energy stored can be easily recovered by allowing the capacitor to discharge in a circuit.

Suppose that, at a given instant, a charge q' has been transferred from one plate to the other. The potential difference V' between the plates at that instant will be q'/C where C is the capacitance.

If an extra amount of charge dq' is to be transferred then, according to eqn. 3.2, the increase in work required will be

$$dW = V'dq' = \frac{q'}{C}dq'$$

The work required to charge the capacitor plates up to a final value q is

$$W = \int dW = \frac{1}{C}\int_0^q q'dq' = \frac{1}{2}\frac{q^2}{C}$$

This work is stored as the potential energy U in the capacitor, so that

$$U = \frac{q^2}{2C} \qquad (4.15)$$

From the relation q = CV, eqn. 4.15 can also be written as

$$U = \frac{C^2V^2}{2C} = \frac{1}{2}CV^2 \qquad (4.16)$$

4.5 Energy density of electric field

It is reasonable to suppose that the energy stored in a capacitor resides in the electric field between its plates just as the energy carried by electromagnetic waves can be regarded as residing in its electric field. As q or V in eqns. 4.15 and 4.16 increases, so does the electric field E; when q or V becomes zero, E vanishes.

In a parallel-plate capacitor, the electric field, neglecting fringing, has the same value for all points between the plates. Thus the energy density U, defined as the stored energy per unit volume, should also be same everywhere between the plates. If A is the area of the capacitor plate and d the distance of separation between the plates, then the volume between the plates is Ad. Hence

$$U = \frac{U}{Ad} = \frac{\frac{1}{2}CV^2}{Ad}$$

Now the capacitance C for a parallel plate capacitor is given by $C = \frac{\epsilon_0 A}{d}$. Eqn. 4.16 thus becomes

$$u = \frac{\frac{1}{2}\frac{\epsilon_0 A}{d}V^2}{Ad} = \frac{1}{2}\epsilon_0\left(\frac{V}{d}\right)^2$$

But V/d is the electric field strength E between the capacitor plates, so that

$$u = \frac{1}{2}\epsilon_0 E^2 \tag{4.17}$$

Although derived for a parallel-plate capacitor, the equations hold no matter what the geometry of the capacitor is. *If an electric field E exists at any point in space (or vacuum), then the point may be regarded as the site of stored energy of* $\frac{1}{2}\epsilon_0 E^2$ *per unit volume.* In general, since E varies with location, u is also a function of coordinates. For the special case of a parallel-plate capacitor, E and U have same value anywhere between the plates.

4.6 Sharing of charges

If two charged capacitors of capacitances C_1 and C_2 be joined by a conductor, some charges will flow from the conductor at higher potential to the conductor at lower potential till the potentials are equalized. Let q_1 and q_2 be the charges on the capacitors after redistribution of charges and let the common potential of the capacitors be V. Then

$$V = \frac{q_1}{C_1} = \frac{q_2}{C_2} = \frac{q_1 + q_2}{C_1 + C_2}$$

Loss of energy due to sharing of charges

When two capacitors of capacitances C_1 and C_2 and potentials V_1 and V_2 are joined by a conductor, electric charges will flow from the capacitor at higher potential to the capacitor at lower potential. Due to this flow of charge, the energy of the system will decrease. The decrease in energy is due to the fact that during the flow of electricity some energy is dissipated in the form of heat.

The total energy of the capacitors before connection

$$= \frac{1}{2} C_1 V_1^2 + \frac{1}{2} C_2 V_2^2$$

The total energy of the capacitors after connection

$$= \frac{1}{2} C_1 V^2 + \frac{1}{2} C_2 V^2$$

where the common potential V is given by

$$V = \frac{C_1V_1 + C_2V_2}{C_1 + C_2}$$

So, the total energy of the capacitors after connection

$$= \frac{1}{2} C_1 \left(\frac{C_1V_1 + C_2V_2}{C_1 + C_2} \right)^2 + \frac{1}{2} C_2 \left(\frac{C_1V_1 + C_2V_2}{C_1 + C_2} \right)^2$$

$$= \frac{1}{2} (C_1 + C_2) \left(\frac{C_1V_1 + C_2V_2}{C_1 + C_2} \right)^2$$

$$= \frac{1}{2} \frac{(C_1V_1 + C_2V_2)^2}{(C_1 + C_2)}$$

Change of energy = final energy – initial energy

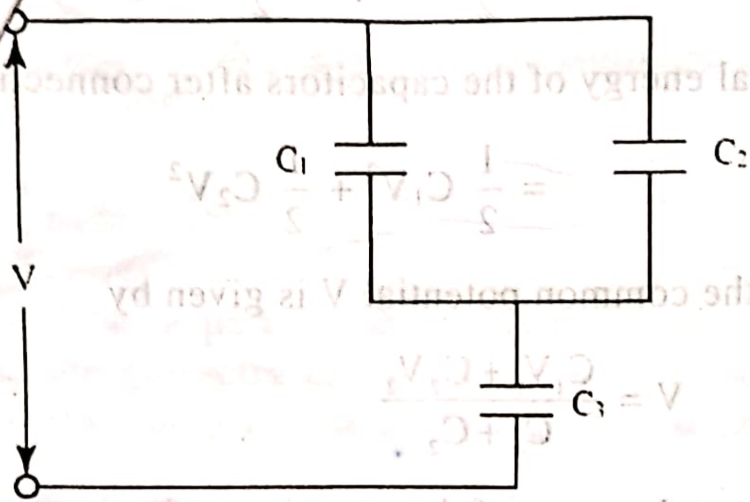
$$= \frac{1}{2} \frac{(C_1V_1 + C_2V_2)^2}{(C_1 + C_2)} - \left(\frac{1}{2} C_1V_1^2 + \frac{1}{2} C_2V_2^2 \right)$$

$$= \frac{1}{2} \left[\frac{(C_1V_1 + C_2V_2)^2 - (C_1 + C_2)(C_1V_1^2 + C_2V_2^3)}{C_1 + C_2} \right]$$

$$= -\frac{1}{2} \frac{C_1C_2(V_1 - V_2)}{C_1 + C_2} \qquad (4.18)$$

The quantity is always a negative quantity. Therefore there will be always a loss of energy when two capacitors are connected by a wire.

Example 4.6 Find the equivalent capacitance of the combination as illustrated in the following figure. Assume $C_1 = 10\mu F$, $C_2 = 5\mu F$, $C_3 = 4\mu F$ and $V = 100$ volts.



Soln.

C_1 and C_2 are connected in parallel. Their equivalent capacitance $C_p = C_1 + C_2 = 10 + 5 = 15\mu F$.

Now, considering C_p connected in series with C_3, we get

$$\frac{1}{C} = \frac{1}{C_p} + \frac{1}{C_3} = \frac{1}{15} + \frac{1}{4} = \frac{4+15}{60} = \frac{19}{60}$$

$$\therefore \ C = \frac{60}{19}\mu F = 3.2\mu F.$$

Example 4.7 Two capacitors have a capacity of $5\mu F$ when connected in parallel and $1.2\mu F$ when connected in series. Calculate their individual capacitances.

Soln.

Let C_1 and C_2 be the individual capacitances when the two are connected in parallel, we have

$$C_1 + C_2 = 5 \qquad \text{(i)}$$

when they are connected in series, we have

$$\frac{1}{C_1} + \frac{1}{C_2} = \frac{1}{1.2}; \qquad \text{or,} \qquad \frac{1}{1.2} = \frac{C_1 + C_2}{C_1 C_2} \qquad \text{(ii)}$$

From (i) and (ii)

$$\frac{1}{1.2} = \frac{5}{C_1C_2}, \quad \text{or,} \quad C_1C_2 = 6 \qquad \text{(iii)}$$

$$\text{Now } (C_1 - C_2)^2 = (C_1 + C_2)^2 - 4C_1C_2$$
$$= 5^2 - 4.6 = 1$$

$\therefore$ $C_1 - C_2 \pm 1$

Let $C_1 - C_2 = +1$

Then, combining with eqn. (i), we get $2C_1 = 6$; or $C_1 = 3\mu F$.

Hence $C_2 = 2\mu F$.

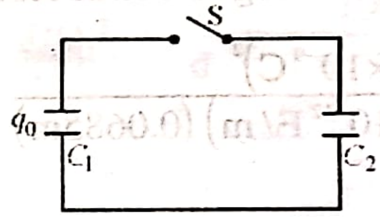
If $C_1 - C_2 = -1$, $2C_1 = 4$ or $C_1 = 2\mu F$ and $C_2 = 3\mu F$.

$\therefore$ $C_1 = 2\mu F$ or $3\mu F$

$C_2 = 3\mu F$ or $2\mu F$.

Example 4.8 *A 3.55μF capacitor C_1 is charged to a potential difference V_0 = 6.30V, using a battery. The charging battery is then removed and the capacitor is connected to an uncharged 8.95μF capacitor C_2 in series – as in the figure. After the switch S is closed, charge flows from C_1 to C_2 until an equilibrium is established, with both capacitors at the same potential difference V. (i) what is the common potential difference V? (ii) what is the energy stored in the electric field before and after the switch S is closed?*

Soln.



(i) when the capacitor C_1 is connected to the capacitor C_2, the original charge q_0 is shared by the two capacitors. Or, $q_0 = q_1 + q_2$.

Applying the relation q = CV, we get

$$C_1V_0 = C_1V_1 + C_2V; \qquad \text{or} \qquad V = \frac{C_1}{C_1 + C_2}V_0$$

$$\text{or, } V = \left(\frac{3.55}{3.55 + 8.95}\right)6.30 = 1.79\text{V}.$$

(ii) The initial stored energy

$$U_i = \frac{1}{2}C_1 V_0^2 = \frac{1}{2}(3.55 \times 10^{-6}F)(6.30V)^2 \simeq 7.05 \times 10^{-5} \text{Joules.}$$

The final energy

$$U_f = \frac{1}{2}C_1V^2 + \frac{1}{2}C_2V^2$$

$$= \frac{1}{2}(3.55 \times 10^{-6}F)(1.79V)^2 + \frac{1}{2}(8.95 \times 10^{-6}F)(1.79V)^2$$

$$= 5.69 \times 10^{-6}J + 14.33 \times 10^{-6}J$$

$$= 20.02 \times 10^{-6}J = 2.0 \times 10^{-6}J.$$

Thus the final energy is less than initial energy by about 72%. This is not a violation of conservation of energy. The 'missing' energy appears as the thermal energy in the connecting wires.

***Example 4.9** An isolated conducting sphere whose radius R is 6.85 cm carries a charge q = 1.25nC. (i) How much energy is stored in the electric field of this charged conductor? (ii) What is the energy density at the surface of the sphere?*

Soln.

(i) The capacitance of an isolated conducting sphere of radius R is $C = 4\pi\epsilon_0 R$.

The energy stored in the electricfield of this charged conductor,

$$U = \frac{q^2}{2C} = \frac{q^2}{8\pi\epsilon_0 R} = \frac{(1.25\times10^{-9}C)^2}{(8)(3.14)(8.85\times10^{-12}F/m)(0.0685m)}$$

$$= 1.03 \times 10^{-7}J = 103nJ.$$

(ii) The energy density,

$$u = \frac{1}{2}\epsilon_0 E^2$$

where $E = \frac{1}{4\pi\epsilon_0} \cdot \frac{q}{R^2}$

$$\therefore \quad u = \frac{1}{2}\epsilon_0 E^2 = \frac{q^2}{32\pi^2 \epsilon_0 R^4}$$

$$= \frac{(1.25\times10^{-9}C)^2}{(32)(3.14)(8.85\times10^{-12}C^2/N.m^2)(0.0685m)^4}$$

$$= 2.54 \times 10^{-5}\ J/m^3 = 25.4\mu J/m^3.$$

***Example 4.10** A plane-parallel plate capacitor has circular plates of radius r = 10.0 cm, separated by a distance d = 1.00 mm. (i) How much charge is stored on each plate when their electric potential difference has the value V = 100V? (ii) Calculate the electric field, the electric field energy density, and the energy stored in the capacitor.*

Soln.

$$A = \pi r^2 = (3.14)\ (0.1m)^2$$

$$d = 1mm = 1.0 \times 10^{-3}m$$

(i) $C = \dfrac{\epsilon_0 A}{d}$

$$= \frac{(8.85\times10^{-12}C^2/N.m^2)(3.14\times10^{-2}m^2)}{1.00\times10^{-3}m}$$

$$= 2.8 \times 10^{-10}F.$$

The magnitude of charge on each plate,

$$q = CV = (2.8 \times 10^{-10}F)\ (100\ V) = 2.8 \times 10^{-8}C = 28nC.$$

(ii) $E = \dfrac{V}{d} = \dfrac{100V}{1.0\times10^{-3}m} = 10^5\ V/m = 100\ kV/m$

energy density, $u = \dfrac{1}{2}\epsilon_0 E^2$

$$= \left(\frac{1}{2}\right)(8.85\times10^{-12}C^2/N.m^2)(100000\ V/m)$$

$$= 0.044\ J/m^3.$$

energy stored, $u = \dfrac{1}{2}CV^2$

$$= \left(\frac{1}{2}\right)(2.8\times10^{-10}\,\text{F})(100\text{V})^2$$

$$= 1.4 \times 10^{-6}\text{J} = 1.4\mu\text{J}.$$

$$[U = u.\pi r^2.d = (4.4 \times 10^{-2}\text{J/m}^3)\,\pi\,(0.1\text{m})^2\,(0.001\text{m}) = 1.4\mu\text{J}]$$

***Example 4.11** Compute the energy stored in a 60 – pF capacitor (i) when charged to a potential difference of 2kV, (ii) when the charge on each plate is 30 nC.*

Soln.

$$\text{(i) } E = \frac{1}{2}CV^2 = \frac{1}{2}(60 \times 10^{-12}\text{F})\,(2000\text{V})^2$$

$$= 1.2 \times 10^{-4}\text{J}.$$

$$\text{(ii) } E = \frac{1}{2}\frac{q^2}{C} = \left(\frac{1}{2}\right)\frac{(30\times10^{-9}\,\text{C})^2}{(60\times10^{-12}\,\text{F})} = 7.5 \times 10^{-6}\text{J}.$$

4.7 Capacitance with dielectric

Up to this point the capacitance was calculated on the assumption that there is no material in the space between the plates of the capacitor. What happens if the space between the plates is filled with a *dielectric*, which is an insulating material such as mineral oil, glass or plastic? The effect of the presence of a dielectric in the space between the plates of a capacitor was first investigated by Michael Faraday.

Faraday constructed two identical capacitors, filling one with dielectric and the other with air under normal conditions. Faraday's experiments showed that when both the capacitors were charged to the *same potential difference, the charge on the capacitor with dielectric was greater than that on the other.* Since q is larger for the same V with the dielectric present, it follows from relation C = q/V that the capacitance of a capacitor *increases* if a dielectric is placed between the plates. If C is the capacitance of the capacitor when a dielectric material is present and C_o the capacitance when no dielectric material is present, then the ratio C/C_o is called the *dielectric constant k* of the dielectric material. Or,

$$k = \frac{C}{C_0}$$

k is a dimensionless factor by which the capacitance of a capacitor increases when a dielectric material is inserted in between the plates, relative to its capacitance when no dielectric is present. It is assumed, unless otherwise stated, that the dielectric completely fills the space between the plates.

The dielectric constant is a fundamental property of the dielectric material and is independent of the size or shape of the conductor. The dielectric constant of some dielectric materials is given in the table below. The dielectric constant of vacuum is unity by definition and for most practical applications, air and vacuum are equivalent in their dielectric effects.

SOME PROPERTIES OF DIELECTRIC MATERIALS

Material	Dielectric constant k	Dielectric strength kV/mm
Vacuum	1 (exact)	∞
Air (1 atm)	1.00059	3
Polystyrene	2.6	24
Paper	3.5	16
Transformer oil	4.5	12
Pyrex	4.7	14
Mica	5.4	160
Porcelain	6.5	4
Silicon	12	
Water (25°C)	78.5	
Water (20°C)	80.4	
Titania ceramic	130	
Strontium titanate	310	8

The values quoted in the table were measured at room temperature.

Some insight into Faraday's experiment is provided by Fig. 4.6. A capacitor is initially charged to a charge q by the battery **B** which then remains connected to the capacitor to ensure that the potential difference V and the electric field E between the plates remain constant. After the dielectric slab is inserted between the plates, the

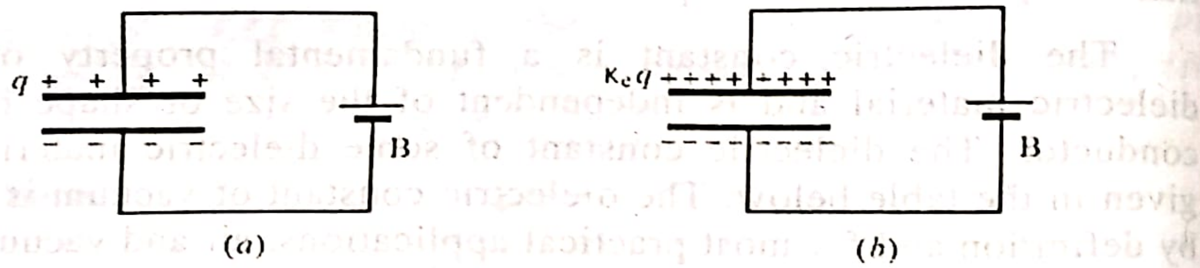


Fig. 4.6

charge increases by a factor k to a value kq. The additional charge $(k - 1)$q is delivered to the capacitor plates by the battery as the dielectric slab is inserted.

Alternately, suppose the battery is disconnected after the capacitor is charged to a charge q as in Fig. 4.7. As the dielectric slab is now inserted, the charge remains constant as there is no path for transfer of charge, but according to relation C = q/V, the potential difference must decrease by a factor k to allow for the increase in capacitance by the same factor. Thus the potential difference decreases by a factor k from V to V/k. Similarly the electric field also decreases from E to E/k.

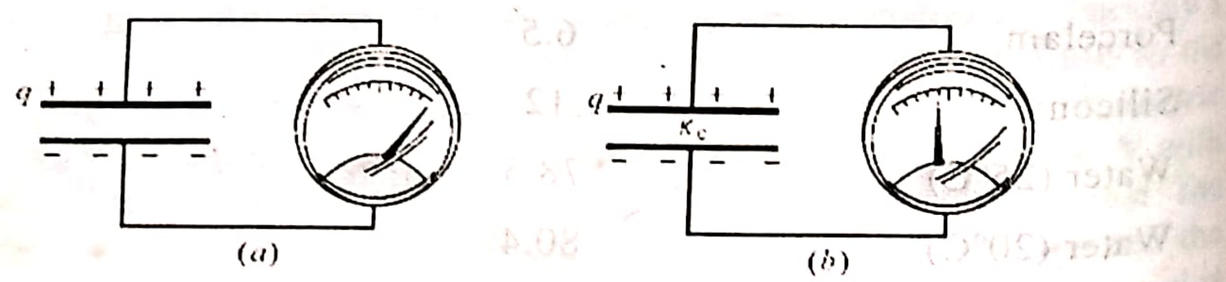


Fig. 4.7

The purpose of a capacitor is to store charge. Thus the presence of a dielectric permits a capacitor to store a factor k more charge for the same potential difference. However, the presence of a dielectric also limits the potential difference that can be maintained across the plates. If this limit is exceeded, the dielectric material breaks down and forms a conducting path between the plates. Every dielectric material has a characteristic *dielectric strength* which is defined as the maximum value of the electric field that it can tolerate without breakdown. A few such values are also listed in the Table.

For a parallel plate capacitor filled with dielectric, the capacitance is given by

$$C = \frac{k\epsilon_0 A}{d} = kC_0$$

where C_0 is the capacitance of the capacitor with air between the plates. The equation suggests that the effect of a dielectric can be summed up in more general terms as

In a region completely filled up by a dielectric, all electrostatic equations containing the permittivity constant ϵ_0 are to be modified by replacing that constant by $k\epsilon_0$.

For a point charge q imbedded in a dielectric, the electric field is given by

$$E = \frac{1}{4\pi k \epsilon_0} \tag{4.18}$$

Eqn. 4.18 gives the total field in the dielectric. As can be seen, the field due to the Coulomb charge is still given by Coulomb's law (without the factor k) but the dielectric itself produces another electric field, which when combined with the field due to the point charge gives eqn. 4.18.

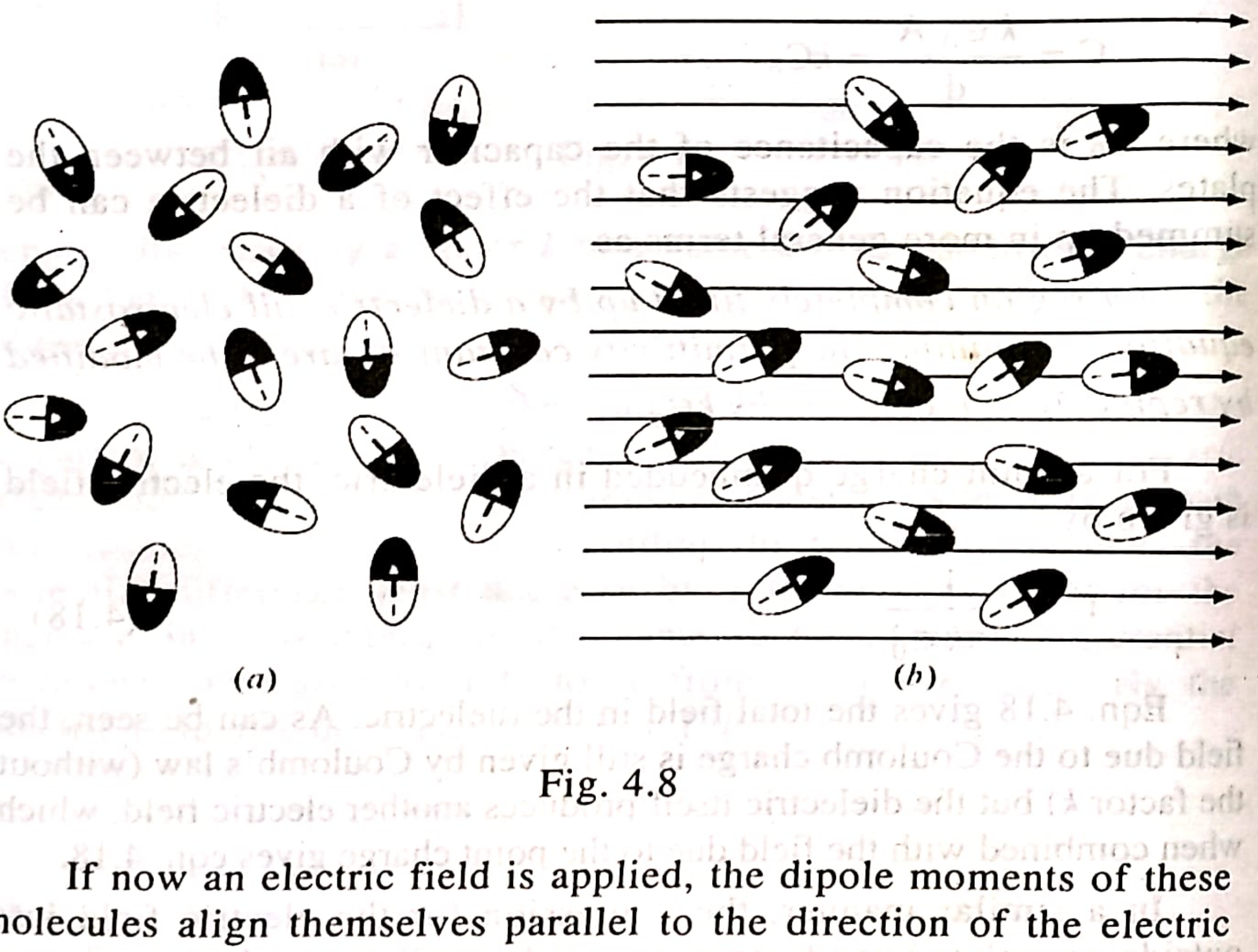
In a similar manner, the expression for the electric field just outside an isolated conductor immersed in a dielectric becomes

$$E = \frac{\sigma}{k\epsilon_0}$$

It is obvious from both of the above expressions that, for a fixed distribution of charges, the effect of a dielectric is to weaken the electric field that would otherwise be present.

4.8 Dielectric: an atomic view

Dielectric materials may be either *polar* or *non-polar.* In the case of molecules of polar dielectrics, *the centres of positive and negative charges do not coincide.* These molecules have permanent electric dipole moments. Some examples of polar molecules are N_2O, H_2O and HCl. In the absence of electric fields, the *electric dipole moments* of these polar molecules are in random orientation (Fig. 4.8) and cancel each other. So even though each molecule has a dipole moment, the average moment per unit volume is zero.



(*a*) (*b*)

Fig. 4.8

If now an electric field is applied, the dipole moments of these molecules align themselves parallel to the direction of the electric field. Because the molecules are in constant thermal agitation, the degree of alignment is not complete [Fig. 4.8(b)] but increases as the applied electric field increases or as the temperature decreases.

In case of non-polar molecules the centres of positive charges and negative charges coincide; so these molecules do not have *any permanent dipole moment.* Oxygen (O_2), nitrogen (N_2) and hydrogen (H_2) molecules are some common examples of non-polar molecules.

2016-17 2014-15

Distinguish between polar and non-polar dielectrics

When a non-polar molecule is placed in an electric field, the centres of positive and negative charges get displaced and the molecules are said to have been *polarized*. Such a molecule is then called *induced electric dipole* and its dipole moment is called *induced electric dipole moment*. The induced electric dipole moment is present only when the electric field is present. It is proportional to the electric field (for normal field strengths) and is created already lined up with the electric field.

Let us consider a slab of a dielectric material as shown in Fig. 4.9. The arrangement of negative and positive charges within the molecules of the dielectric is shown in the figure. This alignment of the dipole moments of the permanent or induced dipoles with the direction of the applied electric field is call *polarization* and the dielectric material and its molecules are said to be *polarized*.

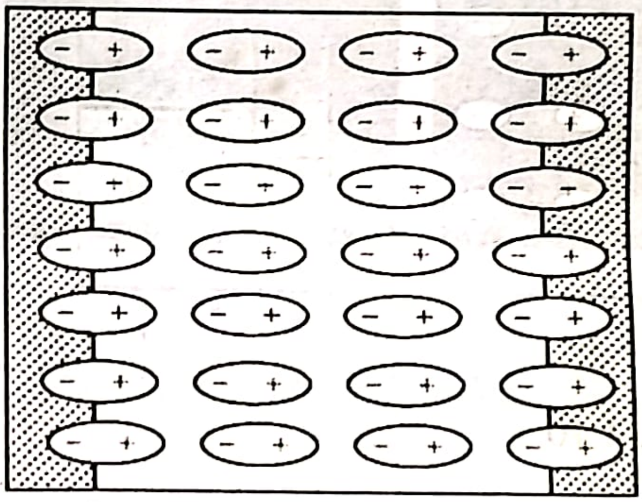


Fig. 4.9

Surface layers extremely thin are shown as shaded part of the figure. Excess positive charge is shown in one thin layer and an equal excess negative charge is shown in the other layer. *The induced charges on the surfaces of the dielectric are due to these layers*. It may be noted that these charges are not free but are bound to the molecules lying in the surface or near the surface. For this reason, these charges are referred to as *bound charges*. Also within the remaining dielectric, the *net charge per unit volume remains zero*. So though the dielectric is polarized, yet it remains *electrically neutral*. Also in polarization, the internal state of the slab is characterized not by an excess charge but by the relative displacement of the charges within it.

Let us use a parallel-plate capacitor, carrying a fixed charge q and not connected to a battery to provide a uniform electric field E_0 into which we place a dielectric slab (Fig. 4.10). The overall effect of alignment and induction is to separate the centre of positive charge of the entire slab slightly from the centre of the negative charge. Although the slab as a whole remains electrically neutral; it has become *polarized.* The net effect is a pile up of positive charge on the right face of the slab and negative charge on the left face, there being no excess charge in any given volume element within the slab. Since the slab as a whole is electrically neutral, the positive induced surface charge must be equal in magnitude to the negative induced surface charge. It should be noted that in this process electrons in the dielectric are displaced from their equilibrium positions by distances that are considerably less than an atomic diameter. There is no transfer of charge over macroscopic distance such as that occurs when a current is set up in the conductor.

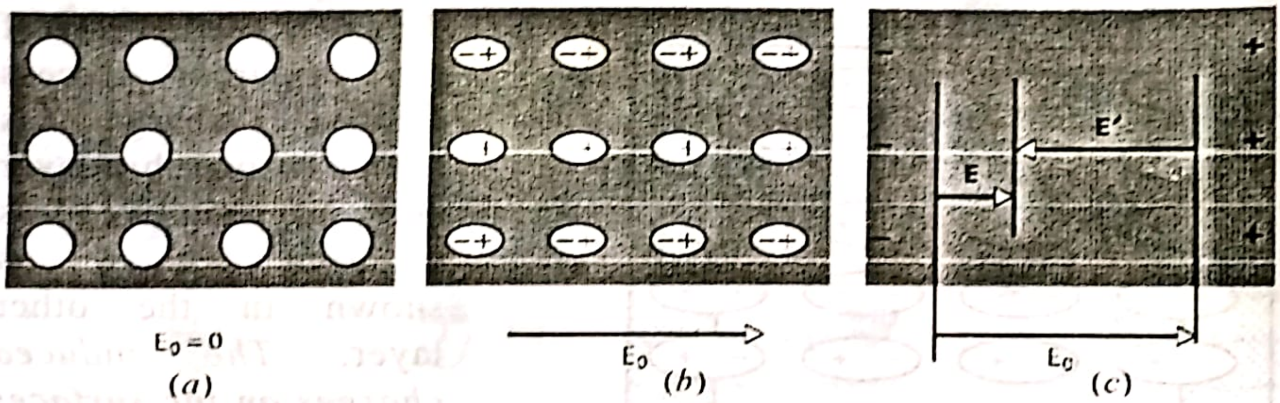


Fig. 4.10

As shown in Fig. 4.10(c), the induced surface charges always appear in such a way that the electric field E' set up by them opposes the external electric field E_0. The resultant field E in the dielectric is the vector sum of E_0 and E'. It points in the same direction as E_0 but is smaller. *If a dielectric is placed in an electric field, induced surface charges appear which tend to weaken original field within the dielectric.*

This weakening of the electric field reveals itself in Fig. 4.7 as a reduction in potential difference between the plates of a charged isolated capacitor when a dielectric is introduced between the plates.

The relation V = Ed for a parallel plate capacitor holds whether or not dielectric is present and shows that the reduction in V described in Fig. 4.7 is directly connected to the reduction in E. More specifically, if a dielectric slab is introduced into a charged parallel-plate capacitor, then

$$\frac{E_0}{E} = \frac{V_0}{V_d} = k$$

4.9 Gauss' law as applied to a dielectric

So far Gauss' law has been confined to situations in which no dielectric is present. We shall now apply the law to a parallel-plate capacitor the space between whose plates is filled with a material of dielectric constant *k*.

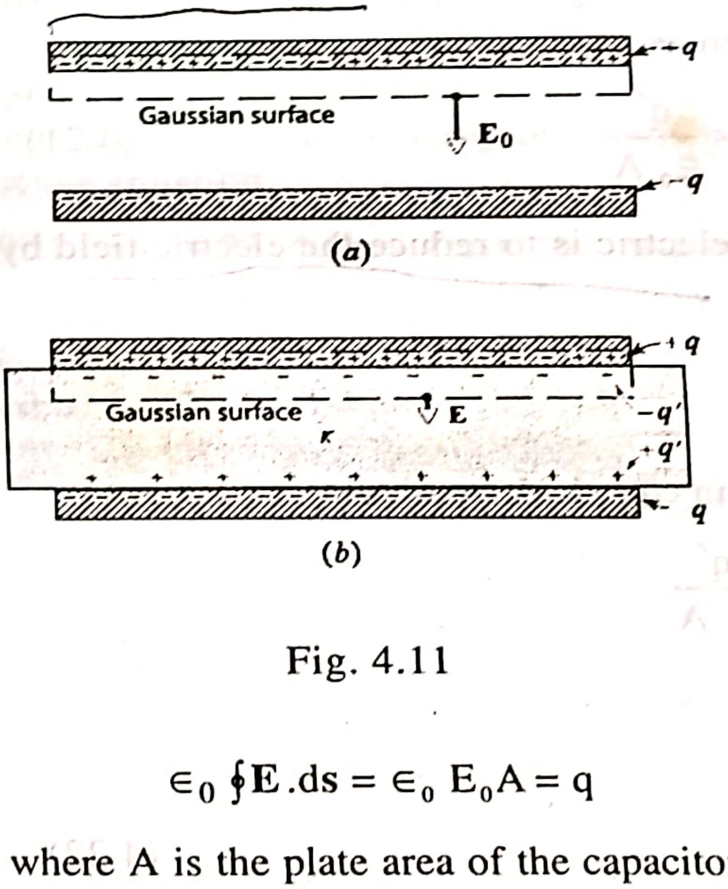


Fig. 4.11

Fig. 4.11 shows a parallel-plate capacitor both with and without the dielectric. The charge q on the plates is assumed to be same in both cases. Gaussian surfaces have been drawn as in Fig. 4.11.

In the absence of the dielectric, Gauss' law gives us

$$\epsilon_0 \oint \mathbf{E}.d\mathbf{s} = \epsilon_0 E_0 A = q$$

where A is the plate area of the capacitor and E_0 is the magnitude of the electric field in the empty space between the plates. E_0 is given by

$$E_0 = \frac{q}{\epsilon_0 A} \tag{4.19}$$

When the empty space between the plates is completely filled up with the dielectric, Gauss' law gives

$$\epsilon_0 \oint E.ds = \epsilon_0 EA = q - q' \qquad (4.20)$$

where E is now the electric field between the plates. The charges q and q' must be distinguished from each other. q is the free charge on the capacitor plates and q' is the induced surface charge which appears on the surfaces of the dielectric under the influence of the electric field E_0, already existing between the capacitor plates at the time of introduction of the dielectric. The induced charge on the dielectric surface adjacent to the positive plate is –q' while that near the negative plate being +q'. Both the charges +q and –q' lie within the Gaussian surface and are opposite in sign. Hence the net charge enclosed within the Gaussian surface is q + (–q') = q – q' as shown in Fig. 4.12.

Fig. 4.20 can be rewritten as

$$E = \frac{q-q'}{\epsilon_0 A} = \frac{q}{\epsilon_0 A} - \frac{q'}{\epsilon_0 A} \qquad (4.21)$$

Now the effect of the dielectric is to reduce the electric field by a factor k. So

$$E = \frac{E_0}{k} = \frac{q}{k\epsilon_0 A}$$

Inserting this value of E in eqn. 4.21 we obtain

$$\frac{q}{k\epsilon_0 A} = \frac{q}{\epsilon_0 A} - \frac{q'}{\epsilon_0 A}$$

$$\text{or, } q - q' = \frac{q}{k}$$

$$\text{or, } q' = q\left(1 - \frac{1}{k}\right) \qquad (4.22)$$

Eqn. 4.22 shows that the induced surface charge q' is always less that the free charge q and is equal to zero if no dielectric is present *i.e.*, *k* = 1 (vacuum).

Substituting the value of q – q' in eqn. 4.20 Gauss' law for the situation when the dielectric is present can be written as

$$\in_0 \oint \mathbf{E}.d\mathbf{S} = \frac{q}{k}$$

Rearranging, we obtain

$$\in_0 \oint k.\mathbf{E}\,.\,d\mathbf{S} = q \qquad (4.23)$$

Eqn. 4.23 is the general form in which Gauss' law is usually written when dielectric is present. Although derived for a parallel-plate capacitor, this important relation is true generally.

The following important points may be noticed regarding eqn. 4.23.

(i) The flux integral now deals with kE, not with E.

(ii) The charge enclosed by the Gaussian surface is taken to be q, *i.e.*, the free charge only. The induced surface charge is deliberately ignored on the right side of the equation, having been taken fully into account by introducing the dielectric constant k on the left side of the equation.

(iii) Eqn. 4.23 differs from the original statement of Gauss' law (eqn. 2.5) only in that $\in_0$ in the latter equation has been replaced by $k\in_0$ in full accord with the statement in Art. 4.7. Moreover, k is taken inside the integral to allow for cases in which k is not constant over the entire Gaussian surface.

Example 4.12 *A parallel-plate capacitor has plates with area A and separation d. A battery charges the plates to a potential difference V_0. The battery is then disconnected and a dielectric slab of thickness d is introduced. Calculate the stored energy both before and after the slab is introduced and account for any difference.*

Soln.

The energy U_0 before the introduction of the slab is

$$U_0 = \frac{1}{2}C_0V_0^2$$

After the introduction of the slab, the energy is

$$U = \frac{1}{2}CV^2$$

where C is the new capacitance and V is the potential difference between the plates.

But $C = kC_0$ and $V = V_0/k$.

where k is the dielectric constant of the slab.

$$\therefore\ U = \frac{1}{2}kC_0\left(\frac{V_0}{k}\right)^2 = \frac{1}{2}\frac{C_0V_0^2}{k} = \frac{1}{k}U_0.$$

Thus the energy after the introduction of the slab is less by a factor $\frac{1}{k}$.

The person who introduces the slab will be aware of this 'missing energy'. He would feel a "tug" on the slab and would have to restrain it if he wants to insert the slab without acceleration. This means that he would have to do negative work on the slab. Alternately, the (condenser + slab) system would do positive work on it.

Example 4.13 A certain parallel-plate capacitor cosists of two plates, each with area 200 cm², separated by a 0.4 cm air gap. (i) Compute its capacitance. (ii) If the capacitor is connected across a 500-V source, what is the charge on it, the energy stored in it, and the value of E between the plates? (iii) If a liquid with k = 2.60 is poured between the plates so as to fill the air gap, how much additional charge will flow onto the capacitor from the 500-V source?

Soln.

(i) For a parallel-plate capacitor

$A = 200\ cm^2 = 0.02 m^2$

$d = 0.4\ cm = 0.004 m$

$$C = \frac{\epsilon_0 A}{d} = \frac{(8.85\times10^{-12})(0.02)}{0.004}$$

$$= 44 \times 10^{-12} F$$

$$= 44\ pF.$$

(ii) $q = CV = (4.4 \times 10^{-11} F)(500V)$

$= 22 \times 10^{-9}C = 22\ nC.$

$$\text{energy} = \frac{1}{2}CV^2 = \left(\frac{1}{2}\right)(44\times10^{-12}F)(500V)^2$$

$= 550 \times 10^{-8}J$

$= 5.5\ mJ.$

$$E = \frac{V}{d} = \frac{500V}{0.004m} = 125 \times 10^3 V/m = 125 kV/m.$$

(iii) The capacitor will now have a capacitance 2.60 times larger than before. Therefore, the charge on the plates of the capacitor will be

$$q' = kCV = (2.6)(44 \times 10^{-12}F)(500V)$$

$= 572 \times 10^{-10}C$

$= 57.2\ nC.$

So, the additional charge that must flow onto it now

$= (57.2 - 22)\ nC = 35.2\ nC.$

***Example 4.14** An isolated conducting sphere whose radius R is 6.85 cm carries a charge q = 1.25nC. (i) How much energy is stored in the electric field of this charged conductor? (ii) What is the energy density at the surface of the sphere? (iii) What is the radius R_0 of a spherical surface such that one-half of the stored potential energy lies within it?*

Soln.

$$U = \frac{q^2}{2C}$$

Now $C = 4\pi\epsilon_0 R$

$$\therefore\ U = \frac{q^2}{8\pi\epsilon_0 R} = \frac{(1.25\times10^{-9}C)^2}{(8)(3.14)(8.85\times10^{-12}F/m)(0.0685m)}$$

$= 1.03 \times 10^{-7}J = 103\ nJ$

(ii) Energy density, $U = \frac{1}{2}\epsilon_0 E^2$

The electric field strength at the surface of the charged conductor is given by

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}$$

$$\therefore \ U = \frac{1}{2}\epsilon_0 E^2 = \frac{q^2}{32\pi^2 \epsilon_0 R^4}$$

$$= \frac{(1.25\times10^{-9}C)^2}{(32)(3.14)^2(8.85\times10^{-12}C^2/N.m^2)(0.0685m)^4}$$

$$= 2.54 \times 10^{-5} \text{ J/m}^3$$

$$= 25.4 \ \mu\text{J/m}^3.$$

(iii)The energy that lies in a spherical shell between radii r and r + dr is given by

$$dU = (4\pi r^2)(dr)(u)$$

where $(4\pi r^2)(dr)$ is the volume of the spherical shell and u is the energy density.

$$\therefore \ dU = \frac{q^2}{32\pi^2 \epsilon_0 r^4} 4\pi r^2 dr = \frac{q^2}{8\pi\epsilon_0}\frac{dr}{r^2}$$

The total energy U is found by integration,

$$\text{or, } U = \int dU = \frac{q^2}{8\pi\epsilon_0}\int_R^\infty \frac{dr}{r^2} = \frac{q^2}{8\pi\epsilon_0 R}$$

[U can also be found from the relation $U = \frac{q^2}{2C}$ where $C = 4\pi\epsilon_0 R$]

Now what will be the radius R_0 of a spherical surface such that half the stored energy lies within it?

In the equation just obtained, let us put

$$\frac{1}{2}U = \frac{q^2}{8\pi\epsilon_0}\int_R^{R_0}\frac{dr}{r^2}$$

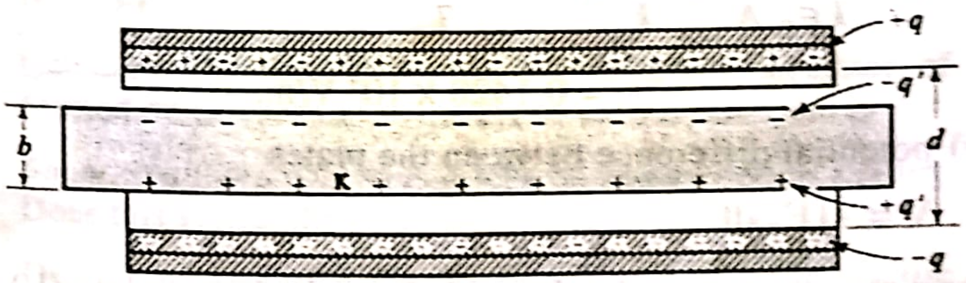
or, $$\frac{q^2}{16\pi\epsilon_0 R} = \frac{q^2}{8\pi\epsilon_0}\left(\frac{1}{R} - \frac{1}{R_0}\right)$$

or, $$\frac{1}{R} - \frac{1}{R_0} = \frac{1}{2R}$$

or, $R_0 = 2R$.

Example 4.15 Fig shows a dielectric slab of thickness b and dielectric constant k placed between the plates of a parallel-plate capacitor of plate area A and separation d. A potential difference V_0 is applied with no dielectric present. The battery is then disconnected and the dielectric slab inserted. Calculate (i) the capacitance C_0 before the slab is inserted, (ii) the free charge q, (iii) the electric field strength in the air gap, (iv) the electric field strength in the dielectric, (v) the potential difference between the plates and (vi) the capacitance with the slab in place.

Assume $A = 100 cm^2$, $d = 1.0$ cm, $b = 0.50$ cm, $k = 7.0$ and $V_0 =$ 100 volts.



Soln.

(i) Capacitance C_0 before insertion of the slab

$$C_0 = \frac{\epsilon_0 A}{d} = \frac{(8.9\times10^{-12}\,coul^2/nt-m^2)(10^{-2}m^2)}{10^{-2}m}$$

$$= 8.9 \times 10^{-12}F = 8.9\ pF.$$

(ii) The free charge q

$$q = C_0V_0 = (8.9 \times 10^{-12}F)\ (100\ V)$$

$$= 8.9 \times 10^{-10} \text{ coul.}$$

(iii)Electric field strength E_0 in the air gap

From Gauss' law for a dielectric, we have

$$\epsilon_0 \oint k\mathbf{E}.d\underline{S} = \epsilon_0\ kE_0A = q$$

for air $k = 1$, hence

$$\epsilon_0 E_0\ A = q$$

$$\text{or, } E_0 = \frac{q}{\epsilon_0\ A} = \frac{8.9\times10^{-10}\text{coul}}{(8.9\times10^{-12}\text{coul}^2/\text{nt}-\text{m}^2)(10^{-2}\text{m}^2)}$$

$$= 1.0 \times 10^4 \text{ volts/m.}$$

(iv)Electric field strength E in the dielectric

From Gauss' law

$$\epsilon_0 \oint \mathbf{E}.\ dS = q$$

$$\text{or, } \epsilon_0 k\,E\,A = q$$

$$\text{or, } E = \frac{q}{k\epsilon_0\ A} = \frac{E_0}{k} = \frac{1.0\times10^4\,V/m}{7}$$

$$= 0.1428 \times 10^4\ V/m.$$

(v) potential difference between the plates

$$V = -\int \mathbf{E}.\ dl$$

For a straight perpendicular path from the lower plate (L) to the upper plate (U),

$$V = -\int_L^U E \cos 180° dl = \int_L^U E\,dl$$

$$= E_0\,(d - b) + Eb$$

$$\text{or, } V = (1.0 \times 10^4 V/m)\ (5 \times 10^{-3} m)$$

$$+ (0.1428 \times 10^4 V/m)\ (5 \times 10^{-3} m)$$

$= 50V + 7.140V = 57.14V.$

(vi) Capacitance C with the slab in place.

$$C = \frac{q}{V} = \frac{8.9\times10^{-10}\text{coul}}{57.14V}$$

$$= 15.57 \times 10^{-12}F$$

$$= 15.57pF.$$

EXERCISES

1. What is a capacitor? Explain the term capacitance. What do you mean by the charge of a capacitor?

2. What guidelines would you follow to calculate the capacitance of a capacitor? Hence calculate the capacitance of a parallel-plate capacitor.

3. Calculate the capacitance of a spherical capacitor of radii a and b. Hence show that the capacitance of an isolated conductor of radius R is $4\pi\epsilon_0 R$.

4. Calculate the capacitance of a cylindrical capacitor of length L formed by two coaxial cylinders of radii a and b.

5. Show that there is always a loss of energy due to sharing of charge. Does this violate the principle of conservation of energy?

6. Derive an expression for the energy stored in the electric field of a charged capacitor. What is energy density? Obtain an expression for it.

7. What is a dielectric material? For a given potential difference does a capacitor store more or less charge with a dielectric than it does without dielectric (vacuum)?

8. A capacitor is charged by using a battery, which is then disconnected. A dielectric slab is then introduced between the plates. Describe qualitatively what happens to the charge, the capacitance, the potential difference, the electric field strength, and the stored energy.

9. A dielectric slab is introduced between the plates of a capacitor while it remains connected to the battery. Describe qualitatively what happens to the charge, the capacitance, the potential difference, the electric field strength, and the stored energy. Does work need to be done to insert the slab?

10. Describe an expression for Gauss' law as applied to a dielectric.

11. Show that Gauss' law as applied to a dielectric is given by

$$\epsilon_0 \oint kE.dS = q$$

where the symbols have their usual meanings.

12. The parallel palates of an air-filled capacitor are 0.5 cm apart. What must be the area of each plate if the capacitance is to be 0.5 farad? **[2.81 x 10^8m^2]**

13. A parallel plate capacitor of capacitance C_1 = 10μF is raised to a potential difference of 150 volts. After disconnecting from the battery it is connected in parallel to a second parallel plate capacitor of capacitance C_2 = 5μF. Find the energy of the system. **[0.075 joules]**

14. A 100 pF capacitor is charged to a potential difference of 50 volts, the charging battery then being disconnected. The capacitor is then connected in parallel to a second capacitor. If the measured potential difference drops to, what is the capacitance of the second capacitor? **[~ 43 pF]**

15. A potential difference of 300 volts is applied to a 2.0μF capacitor. It is then connected to an 8.0μF capacitor in series. What are the charge and potential difference for each capacitor? **[q = 4.8 x 10^{-4}coul, V_1 = 240V and V_2 = 60V]**

16. A capacitor is charged with 9.6nC and has a 120-V potential difference between its terminals. Compute its capacitance and the energy stored in it. **[8.0 x 10^{-11}F, 576 nJ]**

17. Compute the energy stored in a 60-pF capacitor (i) when charged to a potential difference of 2kV, (ii) when the charge on each plate is 30nC. **[1.2 x 10^{-4}J, 7.5 x 10^{-5}J]**

18. Two parallel conducting plates of area 100 cm^2 and 5 mm apart are given equal and opposite charges of 0.20 μc. The region between the plates is filled with a dielectric of k = 5. Compute (i) the capacitance of the system and (ii) the voltage difference between the plates. **[88pF, 2270V]**

19. A parallel plate capacitor of plate area $A = 4 \times 10^{-2} m^2$, plate separation $d = 2 \times 10^{-2} m$ is raised to a potential difference $V_0 = 100$ volts by connecting a battery when there is no dielectric in between the plates. The battery is then disconnected and a dielectric slab of thickness $b = 1 \times 10^{-2} m$ and dielectric constant $k = 5$ is inserted in the middle of the capacitor having equal air spaces near the plates. Calculate (i) the capacitance C_0 before the insertion of the dielectric, (ii) the free charge q, (iii) the field strength E_0 in the air gap, (iv) the field strength in the dielectric, (v) the voltage across the plates after the insertion of the dielectric and (vi) the capacitance after the insertion of the dielectric. **[(i) 17.8pF, (ii) 17.8 x 10^{-10} coul, (iii) 5000 V/m, (iv) 1000 V/m, (v) 60 V and (vi) 29.67pF]**

CHAPTER - V

CURRENT AND RESISTANCE

5.1 Concept of electric current

The previous chapters deal largely with electrostatics, that is, charges at rest. We shall now focus on *electric current*, that is, charges in motion.

When two charged conductors are connected by a metallic wire, electricity flows from the conductor at higher potential to the conductor at lower potential until the two potentials are equal. If the potential difference between the two ends of the wire could be maintained, say by connecting the two ends of the wire to the two terminals of a battery, there would be a steady flow of electricity through the metallic wire. When this happens, we say an electric current is established.

According to free electron theory of metal, the free electrons in an isolated metallic conductor are in random motion like the molecules of a gas confined to a container. In the absence of any applied electric field, electrons have no directed motion along a conductor. If we imagine a hypothetical plane through the conductor, the rate at which electrons cross the plane in one direction is equal to the rate at which they cross in the other direction; the net rate is zero even though there is an abundance of conduction electrons inside the conductor. If the two ends of the conductor is now connected to the two terminals of battery, an electric field **E** is set up everywhere inside the conductor. This electric field **E** acts on the electrons and gives them a net motion in the direction opposite to **E**. If the battery could maintain the potential difference, then the charges (electrons) would continue to circulate indefinitely. This resultant motion of electrons is manifested as electric current. Quantitatively, an electric current is measured by the number of unit of electricity or charges which flow across any cross-section of a conductor per unit time.

If dq is the amount of net charge that passes through any surface in a time interval dt, then the current i is given by

$$i = \frac{dq}{dt} \tag{5.1}$$

or, $dq = i\, dt$

The net charge that passes through the surface in any time interval is found by integrating the current:

$$q = \int i\, dt \tag{5.2}$$

If the rate of flow of charge is constant in time, the current is said to be *steady* or *constant.* Denoting the steady current by I, we simply get

$$I = \frac{q}{t}$$

where q is the total charge that flows in time t. For steady current, charge does not collect at or drain away from any point inside the conductor, *i.e.*, there is no source or sink of charge inside the conductor. Therefore, flow of charge per unit time through any area of cross-section along the conductor is constant, even though the cross-sectional area may be different at different points. Hence, the electric current is the same everywhere along the conductor.

5.2 Direction of the current

The charge carriers in metals are electrons. In electrolytes or gaseous conductors the charge carriers are positive or negative or both. Thus we need a convention for labeling the direction of current because charges of opposite sign move in opposite directions in a given field. A positive charge moving in one direction is equivalent in nearly all external effects to a negative charge moving in the opposite direction. Hence for simplicity and algebraic consistency, the following convention is adopted:

"The direction of the current is the direction in which positive charges would move, even if the actual charge carriers are negative."

In case the charge carriers are negative, they simply move opposite to the direction of the current flow.

Although current is assigned a direction, it is a scalar quantity, not vector. The arrow that is drawn to indicate the direction of the current merely shows the sense of the charge flow through the wire and is *not* to be taken as a vector. As can be seen from Fig. 5.1, current does not obey the laws of vector addition. The current i_1 in

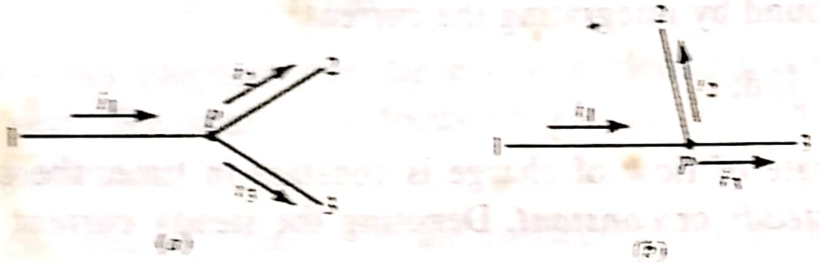


Fig. 5.1

wire 1, divides into two branches i_2 and i_3 in wires 2 and 3, such that $i_1 = i_2 + i_3$. If may be noticed that (a) changing the directions of the wire does not change the way the currents are added – the current in a wire remains unchanged if the wire is bent, tied into a knot, or otherwise distorted and (b) the arrows representing the sense of currents do not in any way obey the laws of vector additions.

The SI unit of current may be defined from eqn. (5.1). It is defined as the quantity of charges in coulombs which flows in one second. This unit is called *ampere* or simply *amp.* (abbreviation A).

Therefore,

amperes = coulombs / seconds

coulombs = ampere – seconds

1 amp – hour = 1 x 60 x 60 = 3600 coulombs.

5.3 Current density

Although steady current I through a conductor is same everywhere, the rate of flow of charge through a unit area of cross-section may be different at different points along the conductor. To describe this flow, a term *current density* is introduced. The current density at a point is defined *as the quantity of charge that flows through a unit area of cross-section around that point per unit time, the unit area being normal to the flow of charges at that point.* It is a vector quantity and is represented by the symbol **j**. It points in the

direction of the electric field and is characteristic of a point inside a conductor rather than of the conductor as a whole.

In electrostatics, the total electric flux through a closed surface s was defined by the equation $\oint_S \mathbf{E}.d\mathbf{A}$ where **E** is the flux density or the electric field. By analogy, the current density **j** for points on any surface-whether planar or not, through which there is a current i, is the electric current flux passing through unit cross-section around that point. Let s be the total surface area through which current i is flowing out and dA is an elementary area around a point where the current density is **j**. Then the total current that escapes the surface area s is

$$i = \int_S \mathbf{j}.d\mathbf{A} \qquad (5.3)$$

where s may or may not be a closed surface.

If the current i is distributed uniformly across a conductor of cross-sectional area A, then j is constant for all points within the conductor and is normal to the cross-sectional area. Then the total current is

$$I = jA$$

$$\text{or, } j = \frac{I}{A} \qquad (5.4)$$

It follows from eqn. (5.4) that the unit of current density is ampere/m^2.

It is also clear from eqn. (5.3) that although **j** is a vector quantity, the electric current i is a scalar quantity since the integral **j.dA** is scalar. Like mass or volume of an object or the length of a rod, the electric current i is a *macroscopic* quantity and is characteristic of the conductor as a whole, whereas **j** is the corresponding *microscopic* quantity and characteristic of a point within the conductor.

5.4 Electron drift velocity

The electric field set up within the conductor by the source of electromotive force, say a battery, exerts a force (= – eE) on the electrons. Hence one may argue that the electrons under the

influence of the electric field should be accelerated as they move along the conductor and thus disturb the steady flow of charge. This, however, is not true. The electric field does not produce a *net* acceleration because the electrons keep colliding with the atoms or ions that make up the conductor. This array of ions, coupled together by strong spring-like forces of electromagnetic origin, is called the *lattice*. The overall effect of the collisions is to transfer kinetic energy from the accelerating electrons into vibrational energy of the lattice. Thus the electrons move through the conductor with a constant average speed, called the *drift speed*, in the direction · **E**, *i.e.*, opposite to the direction of electric field inside the conductor.

The drift speed v_d of the charge carriers in a conductor can be computed from the current density **j**. Fig. 5.2 shows the conduction electrons in a wire moving to the right at an assumed constant drift speed v_d. If n is the number of conduction electrons per unit volume,

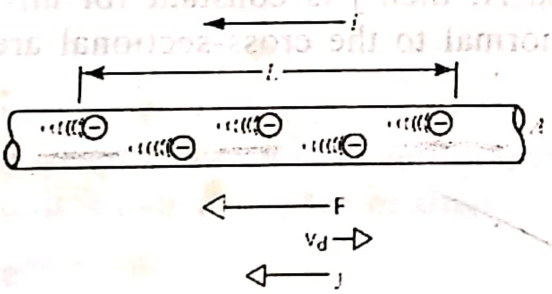


Fig. 5.2

then the number of conduction electrons in a length L of the wire of cross-sectional area A is nAL where AL is the volume of length L of the wire.

The total charge q is, therefore,

$$q = (nAL)\, e$$

where e is the electronic charge. If this charge q flows through the conductor of volume AL in time t, then on the average the electrons cover a distance L in time t. Or,

$$t = \frac{L}{v_d}$$

According to definition of current i, we have

$$i = \frac{q}{t} = \frac{nALe}{L/v_d} = nAev_d$$

$$\text{or,} \qquad v_d = \frac{i}{nAe} \tag{5.5}$$

But j, the current density, is i/A. Hence

$$v_d = \frac{i}{nAe} = \frac{j}{ne} \tag{5.6}$$

$$\text{or,} \qquad j = nev_d$$

Since both j and v_d are vectors, eqn. (5.6) can be written as a vector equation. Following our adopted convention for positive current density, the direction of j must be opposite to $\mathbf{v_d}$. The vector equivalent of eqn. (5.6) is therefore,

$$\mathbf{j} = -\,ne\mathbf{v_d} \tag{7}$$

As can be seen in Fig. 5.2, for electrons, these vectors are indeed in opposite directions.

Example 5.1 *A current of 3.0A flows down a straight metal rod that has a 0.20cm diameter. The rod is 1.5m long and the potential difference between its ends is 40V. Find (a) current density, (b) field in the rod and (c) resistivity of the material of the rod.*

Soln.

(a) $j = \dfrac{I}{A}$ $\qquad\qquad A = \pi r^2 = \pi\left(\dfrac{0.20\times10^{-2}\,m}{2}\right)^2$

$\qquad\qquad\qquad\qquad\quad = \pi \times 10^{-6}\ m^2$

$= \dfrac{3A}{\pi\times10^{-6}\,m^2}$

$= 9.55 \times 10^5\ A/m^2.$

(b) $E = \dfrac{V}{d} = \dfrac{40V}{1.5m} = 26.66\ V/m.$

(c) since $j = E/\rho$ (Art. 5.5)

$$\rho = \frac{E}{j} = \frac{26.66 V/m}{9.55 \times 10^5 A/m^2}$$

$$= 2.79 \times 10^{-5}\ \Omega.m. = 27.9\ \mu\Omega.m.$$

Example 5.2 *A copper wire of 3.0-mm² cross-sectional area carries a current of 5.0A. Find the magnitude of the drift velocity for the electrons in the wire.*

Soln.

$$j = \frac{I}{A} \qquad A = 3.0\ mm^2 = 3.0 \times 10^{-6} m^2$$

$$= \frac{5.0A}{3.0 \times 10^{-6} m^2}$$

$$= 1.67 \times 10^6\ A/m^2.$$

The drift velocity is given by

$$V_d = \frac{j}{ne} = \frac{1.67 \times 10^6 A.m^2}{n(1.60 \times 10^{-19} C)} = \frac{1}{n}(1.04 \times 10^{25}\ m^{-2}.s^{-1})$$

where n is the number of charge carriers per unit volume of copper. Assuming one free electron per atom of copper,

$$n = \frac{(\text{Avogardo's number})(\text{density of copper})}{\text{molecular weight of cooper}}$$

$$= \frac{(6.02 \times 10^{26}\ \text{atoms/kmol})(8920\ kg/m^3)}{63.5\ kg/kmol}$$

$$= 8.5 \times 10^{28}\ \text{atoms}/m^3.$$

$$\therefore\ v_d = \frac{1.04 \times 10^{25} /m^2.s}{8.5 \times 10^{28}\ \text{atoms}/m^3}$$

$$= 0.122 \times 10^{-3}\ m/s$$

$$= 0.122\ mm/s.$$

***Example 5.3** A 0.20 mm diameter copper wire is sealed end to end to a 5.0 mm iron rod, and a current is sent lengthwise through them. If the current in the copper is 8.0A, what are (a) the current and current density in the iron, and (b) current density in the copper?*

Soln.

(a) Since charge must be conserved,

$$I_{Cu} = I_{Fe} = 8.0A.$$

and

$$j_{Fe} = \frac{I}{A}$$

$$= \frac{8.0A}{(3.14)\left(2.5\times10^{-3}\,m\right)^2} \qquad A = \pi r^2 = (3.14)\ (2.5 \times 10^{-3} m)^2$$

$$= 0.40 \times 10^6\ A/m^2$$

$$= 400 \times 10^3\ A/m^2.$$

$$= 400\ kA/m^2.$$

(b) $$j_{Cu} = \frac{I}{A} \qquad A = \pi r^2 = (3.14)\ (10^{-4} m)^2$$

$$= \frac{8.0A}{(3.14)\left(10^{-4}\,m\right)^2}$$

$$= 2.54 \times 10^8\ A/m^2$$

$$= 254\ MA/m^2.$$

***Example 5.4** A strip of silicon, of width $w = 3.2$ mm and thickness $d = 250$ μm, carries a current i of 190 mA. The silicon is an n-type semiconductor, having been 'doped' with a controlled amount of phosphorous impurity. The doping has the effect of greatly increasing n, the number of charge carriers (electrons in this case) per unit volume, as compared with the value for pure*

silicon. In this case, $n = 8.0 \times 10^{21}/m^3$. (a) What is the current density in the strip? (b) What is the drift speed?

Soln.

$$\text{(a) } j = \frac{i}{A} = \frac{i}{wd} = \frac{190\times10^{-3}\,A}{(3.2\times10^{-3}\,m)(250\times10^{-6}\,m)}$$

$$= 2.37 \times 10^5\ A/m^2.$$

$$\text{(b) } v_d = \frac{j}{ne} = \frac{2.37\times10^5\,A/m^2}{(8.0\times10^{21}/m^3)(1.60\times10^{-19}\,C)}$$

$$= 190\ m/s.$$

5.5 Ohm's law, resistivity and conductivity

Let us make a circuit arrangement as shown in Fig. 5.3 and apply a uniform potential difference across the conductor R and measure the resulting current. If the measurement is repeated for various values of the potential difference, it is found that the ratio of

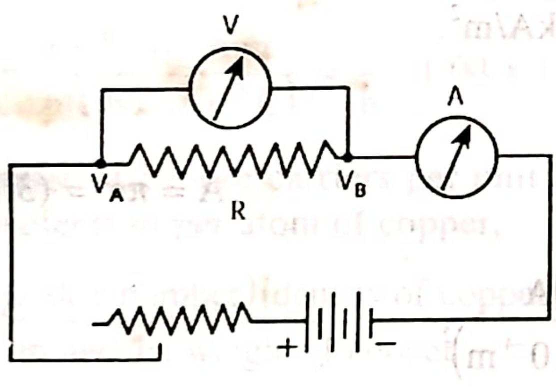


Fig. 5.3

the potential difference to the current is always constant within experimental error, provided the temperature of the conductor is not allowed to change. This result can be summed up by saying that "*the potential difference between the two ends of a conductor always bears a constant ratio with the current flowing through this conductor at a fixed temperature.*" This is known as Ohm's law. Quantitatively Ohm's law can be stated as follows:

$$\frac{V_A - V_B}{I} = \text{constant}$$

where I is the current flowing through the conductor and $V_A - V_B$ is the potential difference between the two ends of a conductor. The plot of $V_A - V_B$ against I will be linear as shown in Fig. 5.4.

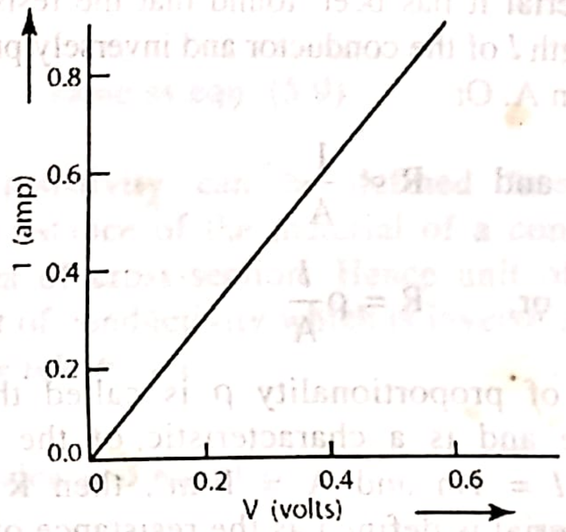


Fig. 5.4

Let another experiment be performed with the same set-up but this time the dimension of the conductor of the same material is increased keeping the length same. It is found that the ratio of the potential difference to the current strength decreases. On the other-hand, if the length of the conductor is increased keeping the dimension same, the ratio increases. This ratio, which is the characteristic of the conductor is called its *resistance* and is usually denoted by R, so that

$$R = \frac{V_A - V_B}{I} = \frac{V}{I} \quad (5.8)$$

where $V = V_A - V_B$ is the potential difference.

If V is in volts and I in amperes, the resistance R is in volts/ampere, which is given the name ohms (abbreviation Ω), so that

1 ohm = 1 volt/ampere

A conductor whose function in a circuit is to provide a specified resistance is called a resistor (symbol –ΛΛΛΛΛΛ–).

Thus, for a fixed potential difference the current through a conductor decreases as R increases. Since the value of R is dependent on the dimension of the conductor, it is purely a characteristic of a given conductor but not a characteristic of the material of the conductor. Like current I, the resistance is a macroscopic quantity. Related to resistance there is another term called the *resistivity*. For a particular conductor of some metallic material it has been found that the resistance is directly proportional to length l of the conductor and inversely proportional to the area of cross-section A. Or

$$R \propto l \quad \text{and} \quad R \propto \frac{1}{A}$$

$$\text{so that } R \propto \frac{l}{A} \quad \text{or,} \quad R = \rho\frac{l}{A} \tag{5.9}$$

where the factor of proportionality ρ is called the *resistivity* or *specific resistance* and is a characteristic of the material of the conductor. When l = 1m and A = 1 m^2, then R = ρ. Thus the resistivity of a material is defined as the resistance of the material of unit length and unit area of cross-section (unit cube of the material).

Resistivity is the microscopic quantity corresponding to resistance. This definition, however, is given for isotropic material. The resistivity is also related to the microscopic quantity **j**, the current density. In an isotropic conductor if **E** is the electric field inside the conductor at a point where the current density is **j**, then **j** is proportional to **E**. Or,

$$\mathbf{j} = \sigma \mathbf{E} \tag{5.10}$$

$$\text{or, } \mathbf{E} = \frac{1}{\sigma}\mathbf{j} = \rho\,\mathbf{j}\,; \qquad \text{or, } \rho = \frac{\mathbf{E}}{\mathbf{j}} \tag{5.11}$$

where σ is called the *conductivity* of the material of the conductor. Inverse of σ is called the resistivity or specific resistance of the conductor.

The relation given by eqn. (5.9) can easily be obtained from the more fundamental eqns. 5.4 $\left(j = \frac{I}{A}\right)$ and 5.10 ($j = \sigma E$). In the case of a conductor of length l and area of cross-section A, the electric field and current density are given by

$$E = \frac{V}{l} \quad \text{and} \quad j = \frac{I}{A}$$

From eqn. 5.11, we have

$$\rho = \frac{E}{j} = \frac{V}{l} \cdot \frac{A}{I} = R\frac{A}{l}$$

$$\text{or, } R = \rho.\frac{l}{A} \quad \text{same as eqn. (5.9)}$$

The unit of resistivity can be defined from eqn. (5.9). Resistivity is the resistance of the material of a conductor of unit length and unit area of cross-section. Hence unit of resistivity is ohm-metre. The unit of conductivity which is inverse of resistivity is then ohm^{-1} $metre^{-1}$ = *mho/metre*.

5.6 Resistors in series and parallel

Just as was the case with capacitors (Chapter. IV), resistors are often used in circuits in various combinations. In analyzing such circuits, it is helpful to replace the combination of resistors by a single resistor, called the *equivalent resistor* R_{eq} whose value should be so chosen that the operation of the circuit remains unchanged.

Resistors connected in parallel

Fig. 5.5 shows two resistors connected in parallel. We would like to find the equivalent resistance of the combination. Let a source of e.m.f. be connected between the points a and b which maintains a potential difference between a and b. The potential difference across

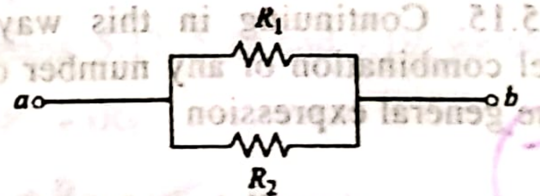


Fig. 5.5

each resistor is the same i.e., V, the potential difference of the source of emf. However, the currents thought the resistors will be different. If i_1 and i_2 are the respective currents through R_1 and R_2, then according to Ohm's law,

$$i_1 = V/R_1 \tag{5.12}$$

and $$i_2 = V/R_2. \tag{5.13}$$

According to the properties of parallel circuit, the total current i must be shared among the branches, so

$$i = i_1 + i_2$$

If the parallel combination is now replaced by a single equivalent resistance R_{eq}, then the same total current i must flow through the circuit. Or,

$$i = V/R_{eq} \tag{5.14}$$

Substituting eqns. 5.12 and 5.13, in 5.14, we obtain

$$\frac{V}{R_{eq}} = \frac{V}{R_1} + \frac{V}{R_2}$$

$$\text{or, } \frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$\text{or, } R_{eq} = \frac{R_1 R_2}{R_1 + R_2} \tag{5.15}$$

If more than two resistors are included in the parallel combination, then the equivalent resistance of the combination may be obtained in the following manner. First find the equivalent resistance R_{12} of the first two resistors R_1 and R_2 using eqn. 5.15. Then find the equivalent resistance of R_{12} and the next resistance R_3, again using eqn. 5.15. Continuing in this way, the equivalent resistance of parallel combination of any number of resistances can be obtained using the general expression

$$\frac{1}{R_{eq}} = \sum_n \frac{1}{R_n} \quad \text{(parallel combination)}$$

$$\frac{1}{R_p} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots\ldots + \frac{1}{R_n} \qquad (5.16)$$

Thus, the reciprocal of the equivalent resistance of a parallel combination is equal to the sum of the reciprocals of the individual resistances of the combination. It may be noted that R_{eq} is always *smaller than* the smallest resistance in the parallel combination. Adding more paths for the current, we get more current for the same potential difference.

Resistances connected in series

Fig. 5.6 shows two resistors connected in series. In series combination of circuit elements, we must travel through *all* the elements in succession.

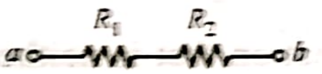


Fig. 5.6

Suppose a battery of potential difference V is connected across the points a and b in Fig. 5.6. A current i is set up in the combination and the same current i passes through each of the resistors. The potential drop across the resistors are

$$V_1 = iR_1 \quad \text{and} \quad V_2 = iR_2 \qquad (5.17)$$

The sum of these potential differences must be equal to the potential difference across the points a and b maintained by the battery, or

$$V = V_1 + V_2 \qquad (5.18)$$

If the combination is replaced by its equivalent resistance, the same current i must flow through the circuit. Hence

$$V = i\,R_{eq} \qquad (5.19)$$

Combining eqns. 5.17, 5.18 and 5.19, we obtain

$$i\,R_{eq} = iR_1 + iR_2$$

or, $R_{eq} = R_1 + R_2$

Extending the result to a series combination of any number of resistors, we obtain

$$R_{eq} = \sum R_n \qquad \text{(series combination)}$$

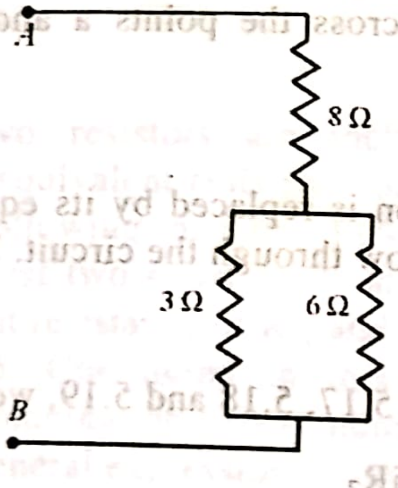
$$= R_1 + R_2 + R_3 + \ldots\ldots\ldots\ldots\ldots\ldots + R_n \qquad (5.20)$$

Thus the equivalent resistance of a series combination is equal to the sum of the individual resistances of the combination. It may be noted that in series combination, the equivalent resistance is always larger than the largest resistance in the combination. Adding more resistors in series means we get less current for the same potential difference.

Comparing the results of series and parallel combination of resistors with those for series and parallel combination of capacitors, it may be noted that resistors in parallel add like capacitors in series, and resistors in series add like capacitors in parallel. This has to do with the different ways resistors and capacitors are defined – resistance being *potential / current* and capacitance being *charge / potential.*

Sometimes, resistors may appear in combination that is neither series not parallel. In such a case the equivalent resistance can be found by breaking the combination into smaller units that can be regarded as series or parallel connections.

Example 5.5 *What is the resistance between A and B in figure below.*



Soln.

For the two resistors in parallel,

$$\frac{1}{R} = \frac{1}{3} + \frac{1}{6} = \frac{2+1}{6} = \frac{3}{6}$$

or, $R = 2\ \Omega$

This is in series with $8\ \Omega$.

So $R_{AB} = 2\ \Omega + 8\ \Omega = 10\ \Omega$.

***Example 5.6** Three resistances of 12, 16 and 20Ω are connected in parallel. What resistance must be connected in series with this combination to give a total resistance of 25Ω ?*

Soln.

The equivalent resistance of the parallel combination is given by

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{12} + \frac{1}{16} + \frac{1}{20} = \frac{20+15+12}{240}$$

$$= \frac{47}{240}$$

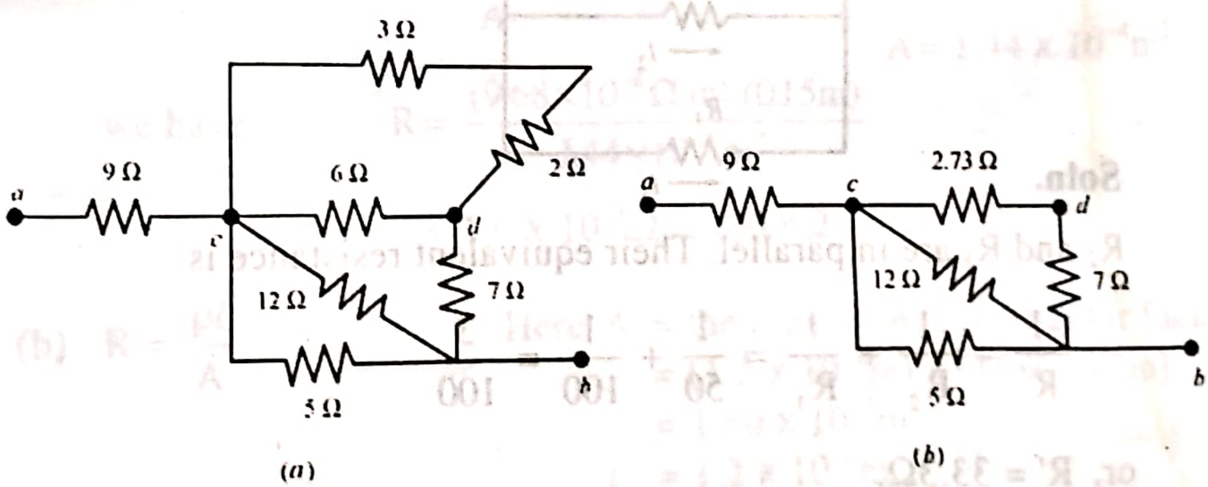
or, $R_{eq} = \frac{240}{47} = 5.11\Omega$.

Then the unknown resistance R_X is obtained from the relation

$R_X + R_{eq} = 25$

or, $R_X = 25 - 5.11 = 19.89\ \Omega$.

***Example 5.7** Find the equivalent resistance between points a and b for the combination shown in figure below.*



Soln.

The 3 and 2Ω resistors are in series and their equivalent resistance is 5-Ω. The 5-Ω resistance is in parallel with 6 Ω resistance. Their equivalent resistance R is given by

$$\frac{1}{R_1} = \frac{1}{5} + \frac{1}{6} = \frac{5+6}{30} = \frac{11}{30}; \text{ or } \quad R_1 = \frac{30}{11} = 2.73\Omega.$$

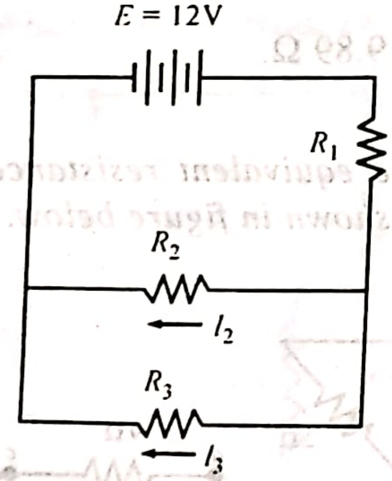
The circuit thus far reduced is shown in (b). The 7 and 2.73- Ω resistances, which are in series, are equivalent to 9.73 Ω. Now the 5, 12 and 9.73 Ω are in parallel and their equivalent resistance is given by

$$\frac{1}{R_2} = \frac{1}{5} + \frac{1}{12} + \frac{1}{9.73} = 0.368, \quad \text{or,} \quad R_2 = 2.6\Omega.$$

This 2.6Ω resistance is in series with 9- Ω resistor. Therefore the equivalent resistance is given by

$$R_{ab} = 9 + 2.6 = 11.6\ \Omega.$$

Example 5.8 The three resistors in the figure below are R_1 = 25Ω, R_2 = 50 Ω and R_3 = 100 Ω. (a) what is the total resistance of the circuit? (b) what are the currents I_1, I_2 and I_3 for a 12-V battery?



Soln.

R_2 and R_3 are in parallel. Their equivalent resistance is

$$\frac{1}{R'} = \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{50} + \frac{1}{100} = \frac{3}{100}$$

or, $R' = 33.3\Omega$.

R' is in series with R_1. Hence the total resistance of the circuit is

$$R = R' + R_1 = 33.3 + 25 = 58.3\Omega.$$

(b) $$I = \frac{E}{R} = \frac{12}{58.3} = 0.206 \text{ A}.$$

The potential across R_2 and R_3 is

$$V' = E - R_1 I = 12V - (25\Omega)(0.206 \text{ A})$$
$$= 6.85 \text{ V}.$$

Therefore,

$$I_2 = \frac{V'}{R_2} = \frac{6.85 \text{ V}}{50 \ \Omega} = 0.137 \text{ A}.$$

$$I_3 = \frac{V'}{R_3} = \frac{6.85 \text{ V}}{100 \ \Omega} = 0.0685 \text{ A}.$$

***Example 5.9** A rectangular block of iron has dimensions 1.2 cm x 1.2 cm x 15 cm. (a) what is the resistance of the block measured between the two square ends? (b) what is the resistance between two opposing rectangular faces? The resistivity of iron at room temperature is 9.68 x 10^{-8} Ω.m.*

Soln.

(a) The area of a square end is $(1.2 \times 10^{-2} \text{m})^2$ or $1.44 \times 10^{-4} \text{m}^2$.

From $R = \frac{\rho l}{A}$ Here $l = 15 \text{ cm} = 0.15\text{m}$

$A = 1.44 \times 10^{-4} \text{m}^2$

we have $$R = \frac{(9.68\times10^{-8}\Omega.\text{m})(0.15\text{m})}{1.44\times10^{-4}\text{m}^2}$$

$$= 1.0 \times 10^{-4}\Omega = 100\text{m}\Omega.$$

(b) $R = \frac{\rho l}{A}$ Here A = the area of the rectangular face

$= (1.2 \times 10^{-2}\text{m}) \times (15 \times 10^{-2}\text{m})$

$= 1.80 \times 10^{-3}\text{m}^2$

$l = 1.2 \times 10^{-2}\text{m}.$

$$= \frac{(9.68\times10^{-8}\,\Omega.\text{m})(1.2\times10^{-2}\,\text{m})}{(1.80\times10^{-3}\,\text{m}^2)}$$

$= 6.5 \times 10^{-7}\Omega = 0.65\mu\ \Omega.$

***Example 5.10** A metal rod of length 20.3 cm has resistance 85 x 10^{-6} ohms. If the diameter of the rod is 0.5cm, calculate the resistivity of the metal.*

Soln.

$$R = \frac{\rho l}{A} \quad \text{or,} \quad \rho = R\frac{A}{l}$$

Here $A = \pi r^2$

$$= (3.14)\left(\frac{0.5}{2}\times10^{-2}\,\text{m}\right)^2$$

$$\rho = \frac{(85\times10^{-6}\,\Omega)(3.14)(0.25\times10^{-2}\,\text{m})^2}{20.3\times10^{-2}\,\text{m}}$$

$l = 20.3$ cm $= 20.3 \times 10^{-2}$m.

$= 0.82 \times 10^{-8}\ \Omega$.m.

***Example 5.11** A nickel wire of length one metre and diameter 0.55 cm has resistance 2.87 x 10^{-3} ohms. If an e.m.f. of one volt is applied between two ends of the wire, calculate (i) current through the wire, (ii) current density and (iii) the electric field strength.*

Soln.

(i) $i = \frac{E}{R} = \frac{1\text{V}}{2.87\times10^{-3}\Omega} = 0.35 \times 10^3$ amp.

(ii) Since the conductor is of uniform cross-section, the current density is everywhere same.

$$\therefore\ j = \frac{i}{A} = \frac{0.35\times10^3\,\text{amp}}{(3.14)(0.275\times10^{-2}\,\text{m})^2}$$

$= 1.47 \times 10^7$ amp/m^2.

Here $A = \pi r^2$

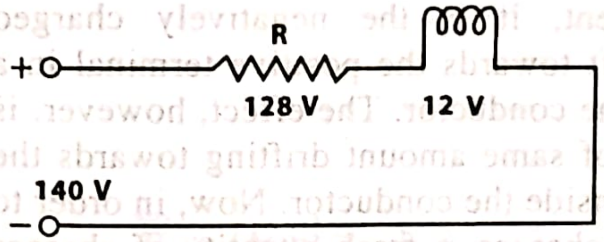
$$= (3.14)\left(\frac{0.55}{2}\times10^{-2}\,\text{m}\right)^2$$

(iii) Electric field strength would be same everywhere inside the wire since current density is everywhere same.

$$E = \frac{1\,\text{V}}{1\,\text{m}} = 1\ \text{volt/m}.$$

***Example 5.12** A 20 watt electric lamp operates at 12 volt D.C. The available D.C. voltage supply is 140 volt. Make a suitable circuit to light the lamp from the supply line using a resistance in series. Calculate the value of the resistance used. What should be the minimum wattage of the resistance to be used?*

Soln.



Since the lamp operates at 12 volt, we can drop the excess voltage, *i.e.*, (140 − 12 = 128) volt across a resistance used in series with the lamp, as in the figure above.

Let I be the current drawn by the lamp.

$$\text{Then } I = \frac{W}{V} = \frac{20W}{12V} = \frac{20}{12} = 1.66A.$$

Since 1.66A must also pass through R, we have

$$R = \frac{V}{I} = \frac{128\ V}{1.66\ A} = 77.1\ \Omega.$$

The power drawn by the resistor R which would appear as Joule's heat, is

$$P = I^2R = (1.66A)^2\ (77.1\Omega) = 212.45W.$$

$$= 0.212\ kW.$$

Hence the minimum wattage of the resistance should be 0.213 kW.

5.7 Electromotive force

We have seen that when a conductor is connected to the two terminals of a battery, a steady current flows through the conductor as long as the potential difference between the ends of the conductor is maintained fixed. Similarly, if an electric appliance is connected to the electric supply line, the electric current drawn by the appliance remains steady. In both cases current is being sustained indefinitely; in the former case by the battery which maintains the potential difference across the two ends of the conductor constant and in the latter case by the supplying

generator. Thus, the battery and the generator are capable of supplying electric energy constantly. The driving energy which sustains the electric current to flow indefinitely is called the *electromotive force* or simply *e.m.f.* (usually denoted by $\mathcal{E}$).

By convention we take the direction of current flow as flow of positive charges from positive terminal to the negative terminal. But in fact, for conduction current, it is the negatively charged conduction electrons which drift towards the positive terminal in a direction opposite to **E** inside the conductor. The effect, however, is equivalent to positive charges of same amount drifting towards the negative terminal *i.e.*, along *E* inside the conductor. Now, in order to sustain the continuous flow of charges, a fresh quantity of charges must be supplied to the positive terminal from the negative terminal. The quantity of charges must be equal to the quantity of charges that leave the positive terminal. Therefore, some work must be done by some external agent to transfer these charges from the lower potential (negative) terminal to the higher potential (positive) terminal. This work is supplied by the battery. *The amount of work done by the battery in bringing a unit positive charge from the negative terminal to the positive terminal is the e.m.f. of the battery.* Thus, if the work done by the battery to transfer dq amount of charge from the negative terminal to the positive terminal is dW, then the electromotive force of the battery is given by

$$\mathcal{E} = \frac{dW}{dq} \tag{5.21}$$

Now let a conductor of resistance R be connected to a battery of *e.m.f.* $\mathcal{E}$ as shown in the Fig. 5.7. Neglecting the resistance of the connecting wires, the potential difference between the ends of the conductor is given, by Ohm's law,

$$V_A - V_B = IR$$

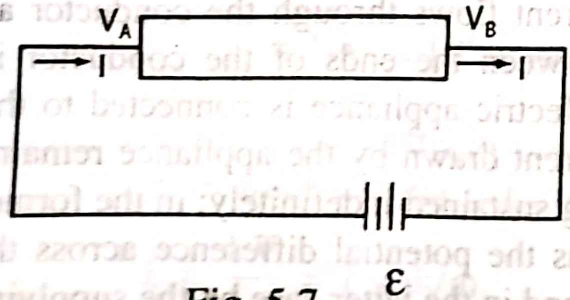
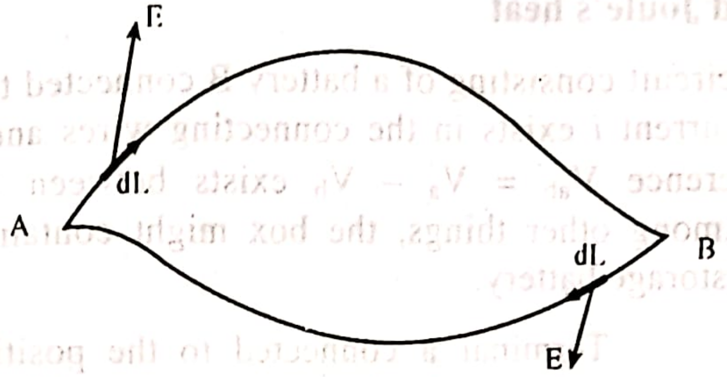


Fig. 5.7

Let us say that a unit charge is moving from point A, corresponding to the high potential term V_A, to the point B of potential V_B (Fig. 5.8). If **E** is the electric field strength



set up by the battery at certain point on the path AB, then the work done in moving a unit charge from A to B is

Fig. 5.8

$$W = \int_A^B \mathbf{E}.dl = -(V_A - V_B)$$

Now if BA is the path along which a unit charge is moving through the battery from the lower potential to the higher potential terminal, then the work done in bringing this charge is

$$\int_A^B \mathbf{E}'.dl = \mathcal{E} \qquad \text{(by definition)}$$

where **E'** is the force on unit charge at some point inside the battery. Since the total work done in a closed path is zero, we have

$$\int \mathbf{E}.\,dl + \int \mathbf{E}'.\,dl = 0$$

$$-(V_A - V_B) + \mathcal{E} = 0$$

Hence $\quad \mathcal{E} = V_A - V_B = IR$

where R is the resistance of the conductor and I is the current flowing through it. Thus the potential difference between the ends of a conductor connected to the terminals of a battery is equal to the *e.m.f.* of the battery itself.

It can be shown in general that if a set of resistors are connected in series to the terminals of a battery, then the *e.m.f.* of the battery is the sum of the potential drop across each of the resistors, *i.e.*,

$$\mathcal{E} = (R_1 + R_2 + R_3 + \ldots\ldots\ldots)\, I$$

5.8 Work, power and Joule's heat

Fig. 5.9 shows a circuit consisting of a battery B connected to a *black box*. A steady current *i* exists in the connecting wires and a steady potential difference $V_{ab} = V_a - V_b$ exists between the terminals a and b. Among other things, the box might contain a resistor, a motor, or a storage battery.

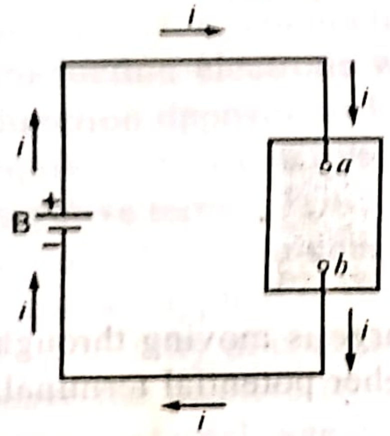


Fig. 5.9

Terminal a connected to the positive terminal of the battery is at a higher potential than terminal b. Let a charge dq flow in a time dt from a to b. The work that must have been done in the process is dqV_{ab}. Therefore, dqV_{ab} is the decrease in potential energy as it moves from a higher potential to a lower potential. From the principle of conservation of energy, this energy is transferred in the box from electrical energy to some other form. What this other form is will depend on what is in the box. If dV is the decrease in potential energy then, the transfer of energy inside the box in time dt is

$$dU = dqV_{ab} = idtV_{ab}$$

The rate of transfer of energy or power P is then given by

$$P = \frac{dU}{dt} = iV_{ab} \qquad (5.22)$$

If the device inside the box is a motor, the energy appears largely as mechanical work done by the motor; if the device is a storage battery that is being charged, the energy appears largely as stored chemical energy in the second battery. If the device is a resistor, the energy appears in the resistor as internal energy associated with an increase of temperature. This is so because, on average, the electrons inside the resistor move with a constant drift speed v_d and hence do not gain any kinetic energy. They lose electric energy through collisions with atoms of the resistor. As a result the amplitudes of the atomic vibrations increase. On a macroscopic scale this corresponds to a temperature increase. Subsequently the energy flows out of the resistor in the form of heat if the environment is at a lower temperature. Thus the loss of energy appears in the form of heat energy.

Substituting the expression for V_{ab} in the expression for P, we obtain

$$P = i\,R\,i$$
$$= i^2 R \tag{5.23}$$

$$\text{or, } P = \left(\frac{V}{R}\right)^2 . R = \frac{V^2}{R} \tag{5.24}$$

Eqn. (5.22) applies to electrical energy transfer of all kinds. Eqns. (5.23) and (5.24) apply only to the transfer of electrical energy to internal energy of a resistor. Eqns. (5.23) and (5.24) are referred to as *Joule's law*, and the corresponding energy transferred to the resistor or its surrounding is called *Joule heating*. Joule's law is a particular type of conservation of energy principle in which the electrical energy in a resistor appears as heat energy.

The unit of power that follows from eqn. (5.22) is volt-ampere. By using the definitions of volt and ampere, the unit of power can be shown to be equivalent to *watt*.

$$1 \text{ volt-ampere} = 1 \frac{\text{joule}}{\text{coulomb}} . \frac{\text{coulomb}}{\text{second}}$$

$$= 1 \frac{\text{joule}}{\text{second}} = 1 \text{ watt.}$$

(non-ohmic) Unilateral circuit

bilateral circuit (ohmic)

5.9 Ohmic and non-ohmic material

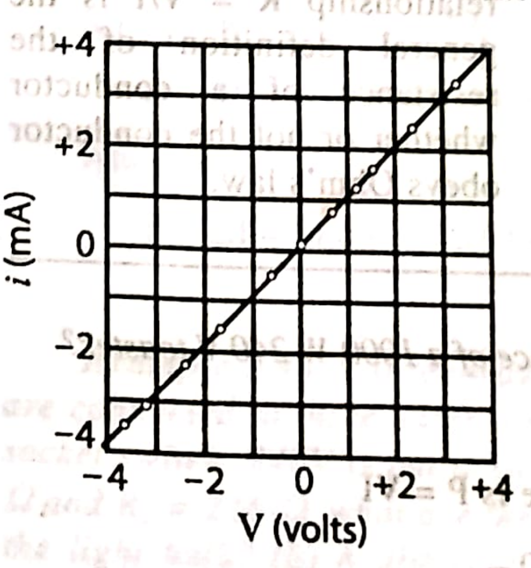


Fig. 5.10

As discussed in Art. 5.5, let us select a particular sample of a conducting material, apply a uniform potential across it, and measure the resulting current. The experiment is repeated for different values of the potential difference. When the resulting current is plotted against the corresponding potential difference, a straight line is obtained (Fig. 5.10). The experimental points clearly fall along the straight line indicating that the ratio V/I is constant. This

means that the resistance of this conductor is constant, independent of the potential across it or the current through it. Further, the line extends to negative potential differences and currents. And we say that the material obeys Ohm's law.

Thus a conducting device is said to obey Ohm's law if the resistance between any pair of points is independent of the *magnitude* and polarity of the *applied potential difference.*

A material or a circuit that obey's Ohm's law is called *ohmic.*

Modern electronic circuits also depend on devices that do not obey ohm's law. The current voltage relationship of a *pn* junction [Fig. 5.11] affords an example for a *non-ohmic* device. It may be noted that the current does not increase linearly with voltage. Furthermore, the device behaves very differently for negative potential differences than it does for positive one.

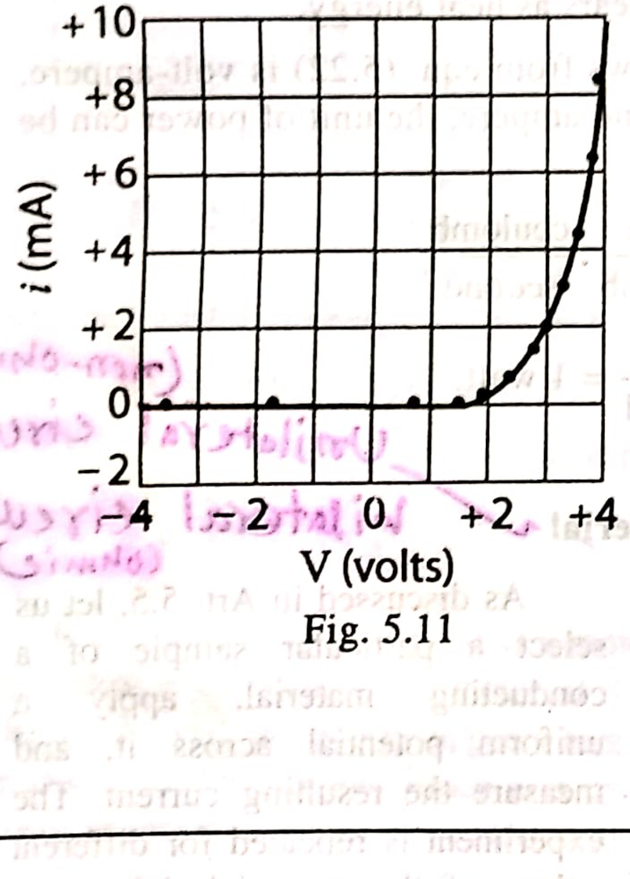


Fig. 5.11

It may be emphasized that the relationship V = iR is *not* a statement of Ohm's law. A conductor obeys Ohm's law only if its V *vs* i graph is linear, *i.e.*, if R is independent of V and i. The relationship R = V/I is the general definition of the resistance of a conductor whether or not the conductor obeys Ohm's law.

***Example 5.13** What is the resistance of a 1000-W 240 V toaster?*

Soln.

Power expended in the resistance is P = VI

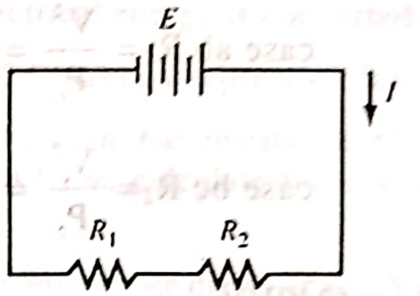
or 1000 = 240I, or $I = \frac{1000}{240} = 4.16$ A.

From Ohm's law, $V = IR$, or $R = \frac{V}{I} = \frac{240}{4.16} = 57.6\Omega$

Alternately,

$$P = \frac{V^2}{R}, \quad \text{or } R = \frac{V^2}{P} = \frac{(240)^2}{1000} = 57.6\Omega.$$

***Example 5.14** Suppose the emf of the ideal battery in the figure below is 4.5V and the resistors are $R_1 = 3\Omega$ and $R_2 = 6\Omega$. What are (a) the total resistance of the circuit, (b) the current in the circuit, and (c) the power dissipated in each resistor?*



Soln.

(a) $R = R_1 + R_2 = 3\Omega + 6\Omega = 9\Omega$

(b) $I = \frac{E}{R} = \frac{4.5V}{9\Omega} = 0.5A$

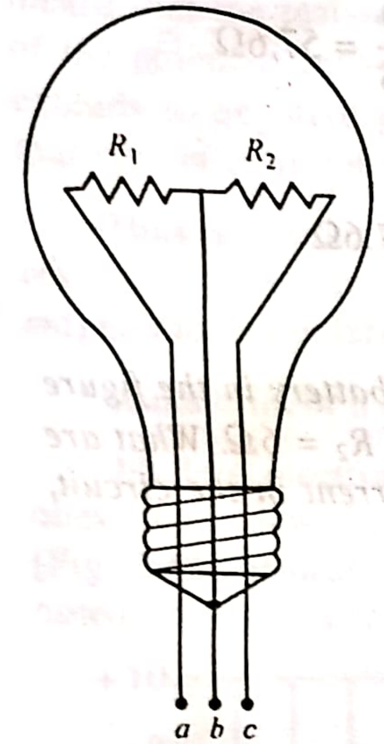
(c) $P_1 = R_1I^2 = (3\Omega)(0.5A)^2 = 0.75\ W$

$P_2 = R_2I^2 = (6\Omega)(0.5A)^2 = 1.50\ W$

Alternately,

$P = EI = (4.5V)(0.5A) = 2.25W = P_1 + P_2$

***Example 5.15** A three-way light bulb has two filaments which are connected to three wires, as shown in the figure. By turning the socket switch, 240V is put across either ab, bc or ac (a) If $R_1 = 144\ \Omega$ and $R_2 = 216\ \Omega$, what are the three possible power dissipations of the light bulb? (b) A different three-way light bulb can operate at 300, 100 and 75 W. What are the resistances of its two filaments?*



(a) case ab $P_1 = \dfrac{V^2}{R_1} = \dfrac{(240V)^2}{144\,\Omega} = 400$ W

case bc $P_2 = \dfrac{V^2}{R_2} = \dfrac{(240V)^2}{216\,\Omega} = 267$ W

case ac $P_3 = \dfrac{V^2}{R_1 + R_2} = \dfrac{(240V)^2}{(144+216)\,\Omega} = 160$ W.

(b) The highest wattage occurs for the smallest resistance, so

case ab $R_1 = \dfrac{V^2}{P_1} = \dfrac{(240V)^2}{300W} = 192\,\Omega$

case bc $R_2 = \dfrac{V^2}{P_2} = \dfrac{(240V)^2}{100W} = 576\,\Omega$

case ac $R_3 = \dfrac{V^2}{P_3} = \dfrac{(240V)^2}{75W} = 768\,\Omega = R_1 + R_2.$

Example 5.16 *You are given a length of a nichrome wire having a resistance R = 72 ohm. It has to be connected across a 120-V line. Under which circumstances will the wire dissipate more heat (a) its entire length is connected across the line, or (b) the wire is cut in half and the two halves are connected in parallel across the line?*

Soln.

(a) The power P dissipated by the entire wire is given by

$$P = \frac{V^2}{R} = \frac{(120V)^2}{72\,\Omega} = 200 \text{ W}.$$

(b) The power for a wire of half-length (and thus half resistance) is

$$P' = \frac{V^2}{\frac{1}{2}R} = \frac{(120V)^2}{36\,\Omega} = 400 \text{ W}.$$

There are two halves so that the power obtained from both of them is 800W, or four times that for a single wire.

EXERCISES

1. Explain the terms current density and drift velocity of electrons inside a conductor. Show that the two are related by the relation $\mathbf{j} = -ne\mathbf{v_d}$.

2. State and explain Ohm's law. What do you mean by resistance of a conductor?

3. Distinguish between resistance and resistivity. What are ohmic and non-ohmic materials?

4. Prove the series and parallel laws of combination of resistances.

5. What is electromotive force? When a source of electromotive force maintains a steady current through a resistance in a circuit, obtain an expression for the rate at which electrical energy is converted into heat energy.

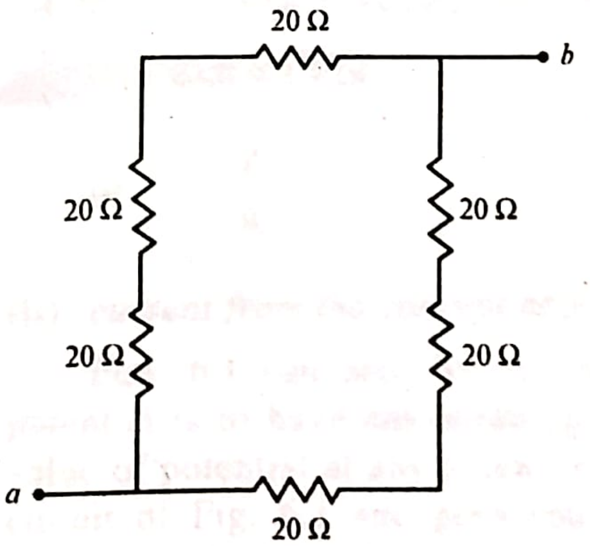
6. What is Joule's heating? Obtain an expression for Joule's heat.

7. A metal rod of length 20.3 cm. has resistance 85×10^{-6} ohms. If the diameter of the rod is 0.5 cm, calculate the resistivity of the metal. **[0.82×10^{-8} Ω-m]**

8. One end of an aluminium wire whose diameter is 2.5 mm is welded to one end of a copper wire whose diameter is 1.8 mm. The composite wire carries a steady current i of 1.3A. (i) What is the current density in each wire? (ii) What is the drift speed of conduction electrons in the copper wire. Take the number of electrons per unit volume in the copper wire to be 8.49×10^{28} electrons/m^3. **[(i)26A/cm^2, 51A/cm^2 (ii) 3.8×10^{-5} m/s]**

9. A copper wire is 20m long and 0.254 mm in diameter. Compute its resistance. Resistivity of copper is 1.7×10^{-8} Ω.m. **[6.7 Ω]**

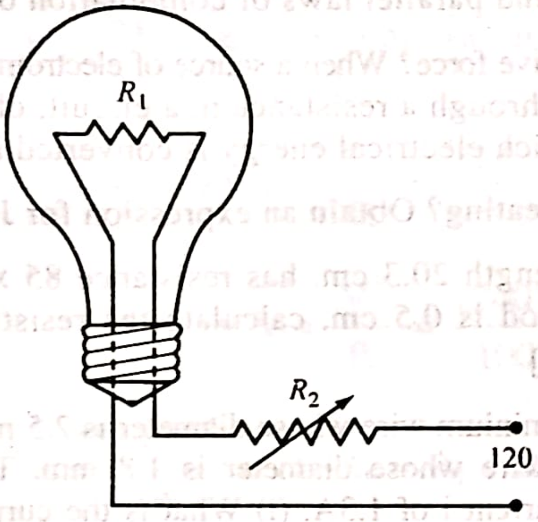
10. What is the resistance between a and b in the figure below.



[R = 30 Ω]

11. An ideal battery with $\mathcal{E}$ = 12V is connected in series to two bulbs with resistances R_1 = 2Ω and R_2 = 4Ω. What is the current in the circuit and the power dissipation in each bulb? **[2A, 8W, 16W]**

12. Fig. shows a potential difference of 120V placed across a circuit that has a lamp with resistance R_1 = 144 Ω connected in series to a variable resistor R_2. The brightness of the lamp is controlled by changing the magnitude of R_2. Find the power dissipations in the lamp (a) when R_2 is zero and (b) when R_2 = 144Ω. (c) what must R_2 be for the power dissipation in the lamp to be 50W? **[100W, 25W, 203.7Ω]**



13. An ideal battery dissipates 2.6W when connected to a 125-Ω resistor. What are (a) the *emf* of the battery and (b) the current in the resistor? **[$\mathcal{E}$ = 18V, I = 0.144 A]**

CHAPTER - VI

D. C. CIRCUITS

6.1 Calculation of current in a single loop

Let us consider a single loop circuit, as that shown in Fig. 6.1, containing one source of emf $\mathcal{E}$ and one resistor R. We would like to find out the current i flowing through the circuit.

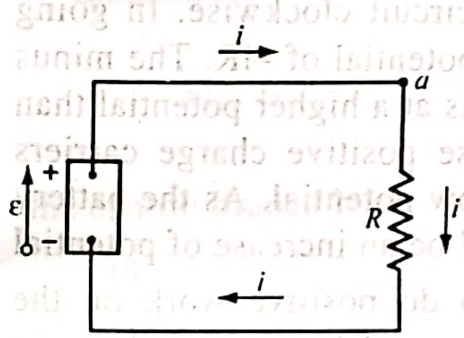


Fig. 6.1

There are two ways of finding the current in a single loop circuit, one based on the principle of conservation of energy and the other on the concept of potential. The two are completely equivalent because potential difference are defined in terms of work and energy.

(i) *current from principle of conservation of energy*

If the current in the circuit is i, then in a time dt an amount of energy given by i^2 Rdt appears in the resistor as internal energy. During this same time a charge dq (= idt) moves through the seat of emf. The work done by the seat of emf to move this charge is given by

$$dW = \mathcal{E}dq = \mathcal{E}idt$$

From the principle of conservation of energy, the work done by the seat must equal the internal energy deposited in the resistor, or

$$\mathcal{E}idt = i^2Rdt$$

$$\text{or,} \quad i = \frac{\mathcal{E}}{R} \qquad (6.1)$$

(ii) *current from the concept of potential*

Eqn. 6.1 can also be derived by considering that if electric potential is to have any meaning, a given point can have only one value of potential at any given time. If one starts at any point in the circuit of Fig. 6.1 and goes round the circuit in either direction, adding up algebraically the changes in potential that one encounters,

then one must encounter the same potential upon return at the starting point. This rule can be summarized as follows:

The algebraic sum of the changes in potential encountered in a complete traversal of any closed circuit is zero.

The above statement is referred to as Kirchhoff's *second law* – for brevity referred to as *loop rule.*

Referring to Fig. 6.1 let us consider a starting point, say a, whose potential is V_a, and traverse the circuit clockwise. In going through the resistor there is a change of potential of $-iR$. The minus sign indicates that the top of the resistor is at a higher potential than the bottom, which must be true, because positive charge carriers move of their own accord from high to low potential. As the battery is traversed form bottom to top, there will be an increase of potential equal to $+\mathcal{E}$ because the battery has to do positive work on the charge carriers in moving them from a point of low potential to one of high potential. When the changes in potential are added algebraically to the initial potential V_a, the sum must yield the final value which must be identical to V_a. Or.

$$V_a - iR + \mathcal{E} = V_a$$

This can be written in the form

$$-iR + \mathcal{E} = 0$$

which is independent of V_a. This explicitly asserts that the algebraic sum of the potential changes for the traversal of a complete circuit is zero. This relation leads to

$$i = \mathcal{E}/R$$

which is identical with eqn. 6.1

6.2 Multi-loop circuits – Kirchhoff's laws

When circuits containing more than one loop are analyzed, it is helpful to consider their *junctions* and *branches.* A junction in a multi-loop circuit is a point in the circuit at which three or more wire segments meet. A multi-loop circuit is shown in Fig. 6.2. For

simplicity, the internal resistances of the batteries have been neglected. There are two junctions in this circuit, at b and d. Points a and c are *not* junctions, because only two wire segments meet at those point.

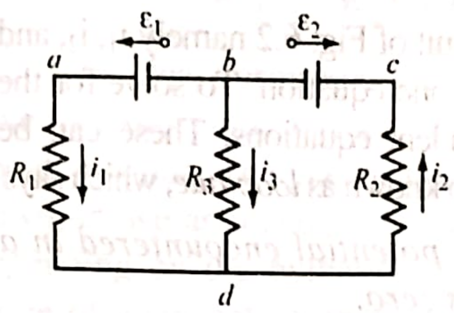


Fig. 6.2

A branch is any circuit path that starts on one junction and proceeds along the circuit to next junction. There are three branches in the circuit of Fig. 6.2, that is, there are three paths that connects junctions b and d – the left branch bad, the right branch bcd and the central branch bd.

In single - loop circuits, there is only one current to determine. In multi-loop circuits, however, each branch has its own individual current, which must be determined in the analysis of the circuit. The three currents in the three branches are labeled i_1 (branch bad), i_2 (branch bcd) and i_3 (branch bd). The directions of the currents have been chosen arbitrarily. Looking at the figure carefully, it may be noted that i_3 must point in a opposite direction to the one shown in the figure. This has been done deliberately to show how the formal mathematical procedures always correct such incorrect guesses.

Flowing of current means carrying of charge. At junction d, the total rate at which charge enters the junction is given by i_1+i_3, and the rate at which charge leaves is given by i_2. In the steady state condition of the circuit, charge does not collect at or drain away form any point in the circuit, the rate at which charge enters a junction must be equal to the rate at which it leaves the junction. We therefore, have

$$i_1 + i_3 = i_2 \qquad (6.2)$$

$$(\text{or, } i_1 + i_3 - i_2 = 0)$$

Eqn. 6.2 suggests a general principle for the solution of multi-loop circuits:

The sum of the currents arriving at any junction must be equal to the sum of the currents leaving that junction

OR

The algebraic sum of the currents meeting at a junction is zero.

This rule is often called *Kirchhoff's first law* and is also referred to as *junction rule.*

Now, there are three unknowns in the circuit of Fig. 6.2 namely i_1, i_2, and i_3. Applying Kirchhoff's first law we get only one equation. To solve for the three unknowns, we need two more independent equations. These can be obtained by applying *Kirchhoff's second law*, also known as *loop rule*, which says

The algebraic sum of the changes in potential encountered in a complete traversal of any closed circuit is zero.

OR

The algebraic sum of the products of currents and resistances in each of the conductors in any closed path (or mesh) in a network plus the algebraic sum of the emfs in that path is zero $[\Sigma iR + \Sigma \varepsilon = 0]$

In order to find the potential differences, the following rules will be observed:

(i) If a resistance is traversed in the direction of the current, the change in potential is $-iR$; in the opposite direction it is $+iR$.

(ii) If a seat of emf is traversed in the direction of the emf *i.e.*, form the negative terminal to the positive terminal, the change in potential is $+\varepsilon$; in the opposite direction it is $-\varepsilon$.

If we traverse the left loop of Fig. 6.2 in a counter clockwise direction starting and ending at point b, the loop rule gives

$$\varepsilon_1 - i_1R_1 + i_3R_3 = 0 \tag{6.3}$$

It is immaterial whether the loop is traversed in a clockwise or counterclockwise direction. The resulting equation will be same.

Traversing the right loop, again in a counterclockwise direction gives

$$-i_3R_3 - i_2R_2 - \varepsilon_2 = 0 \tag{6.4}$$

These two equations, together with eqn. 6.2 derived earlier on the basis of the junction rule, are the three simultaneous equation needed to solve for the unknowns i_1, i_2 and i_3. Solving we get

$$i_1 = \frac{\varepsilon_1(R_1 + R_2) - \varepsilon_2 R_3}{R_1R_2 + R_2R_3 + R_1R_3}$$

$$i_2 = \frac{\varepsilon_1 R_3 - \varepsilon_2(R_1 + R_3)}{R_1R_2 + R_2R_3 + R_1R_3} \tag{6.5}$$

$$i_3 = \frac{-\mathcal{E}_1 R_2 - \mathcal{E}_2 R_1}{R_1 R_2 + R_2 R_3 + R_1 R_3}$$

If is clear form eqn. 6.5 that no matter what numerical values are given to the emfs and to the resistances, the current i_3 always has a negative value. This means that it always points up in Fig. 6.2 rather then down, as we assumed. The currents i_1 and i_2 may be in either direction, depending on the numerical values of the emfs and resistances. If the value of a current so obtained is +ve, the direction of the current is the same as indicated arbitrarily, but if –ve, the direction of the unknown current is opposite to that shown arbitrarily.

6.3 Application of Kirchhoff's Laws

When conductors are connected in a manner, so that they can not be reduced to simple series and parallel arrangement, *i.e.* when we are dealing with complicated electrical circuit, it is possible to determine the combined resistances, or the current in any part of the circuit, by using Kirchhoff's laws. One such complicated circuit which we usually come across in experimental work, is the **Wheatstone bridge** or **net**. This bridge consists of four resistances of values P, Q, R and S ohms connected to a cell E of internal resistance b and emf E volts, and a galvanometer of resistance G ohms, as shown in Fig. 6.3. Suppose we want to find the current i_g through the galvanometer at any instant when the bridge is unbalanced. The entire arrangement can be divided into three networks or meshes (1), (2) and (3). The values of the currents through the galvanometer and through the various branches can be calculated.

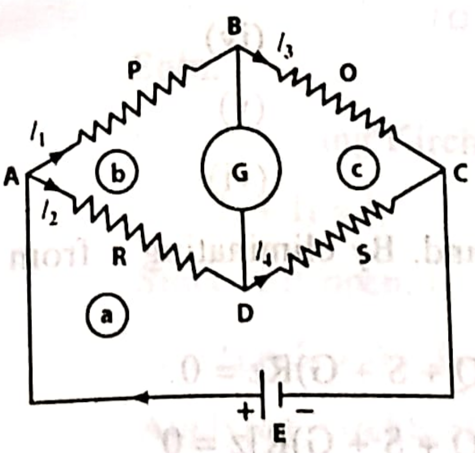


Fig. 6.3

or

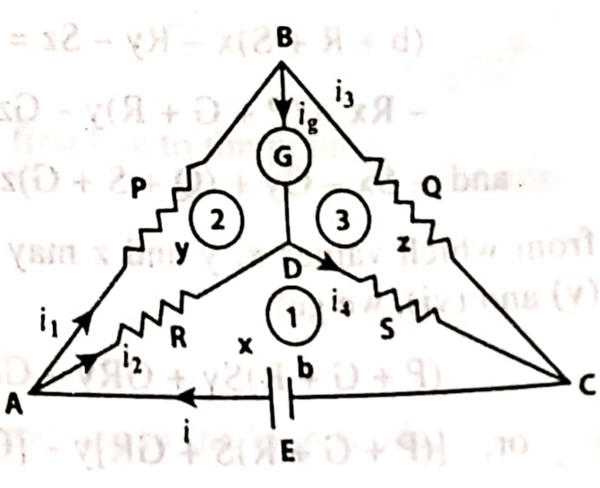


Fig. 6.4

The calculation of the current through the galvanometer or any other branch of the network is very complicated even by the direct application of Kirchhoff's laws. Maxwell, therefore, suggested the following modification which make the calculation more simple:

In each mesh a cyclic current of some specific value is imagined to flow – all these cyclic currents being in the same direction. The emf encountered in traversing the current is reckoned positive if it sends a current in the assumed direction of the mesh current, but negative if it tends to send the current in the opposite direction.

The source of the assumed current is usually taken as *clockwise*. If on solution of the problem the current in any part of the circuit is found to be negative; the current actually flows in a counter clockwise direction.

Let us assume that currents of values x, y, z flow in the loops or meshes (1), (2) and (3) as shown in Fig. 6.4.

For mesh (1), we have

$$bx + R(x - y) + S(x - z) = E \qquad \text{(i)}$$

For mesh (2), we have

$$Py + G(y - z) + R(y - x) = 0 \qquad \text{(ii)}$$

And for mesh (3), we have

$$Qz + S(z - x) + G(z - y) = 0 \qquad \text{(iii)}$$

Relations (i), (ii) and (iii) can be rearranged as

$$(b + R + S)x - Ry - Sz = E \qquad \text{(iv)}$$

$$-Rx + (P + G + R)y - Gz = 0 \qquad \text{(v)}$$

and $-Sx - Gy + (Q + S + G)z = 0$ (vi)

from which values x, y and z may be found. By eliminating x from (v) and (vi), we get

$$(P + G + R)Sy + GRy - GSz - (Q + S + G)Rz = 0$$

or, $[(P + G + R)S + GR]y - [GS + (Q + S + G)R]z = 0$

For no current in the galvanometer y = z

$$\therefore \quad (P + G + R)S + GR = GS + (Q + S + G)R$$

or, $\quad PS = QR$

or, $$\frac{P}{Q} = \frac{R}{S} \tag{6.6}$$

which is the Wheatstone's principle.

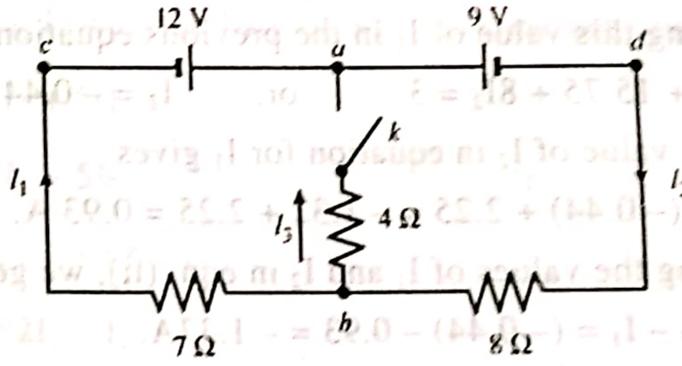
When, however, the bridge is off balance, the current through the galvanometer is either y − z or z − y, and may be easily found in terms of P, Q, R and S.

If R_t = joint resistance of the net-work, the total resistance = $R_t + b$.

Hence by Ohm's law, we get

$$x = \frac{E}{R_t + b}; \qquad R_t = \frac{E}{x} - b.$$

Example 6.1 Find I_1, I_2 and I_3 in the circuit given below (i) when the switch K is open and (ii) when the switch is closed.



Soln.

(i) Applying Kirchhoff's first law to the point a,

$I_1 + I_3 = I_2$

Since K is open, $I_3 = 0$

$\therefore \quad I_2 = I_1 + 0 = I_1$

Applying Kirchhoff's second law to the loop acbda,

$$-12 + 7I_1 + 8I_2 + 9 = 0 \qquad \text{(i)}$$

or, $15I_1 = 3 \quad (\because I_1 = I_2)$

or, $I_1 = \frac{3}{15} = 0.2$ A.

(ii) when the switch K is closed, I_3 is no longer zero. Applying the first law to the point a, we get

$$I_1 + I_3 = I_2 \qquad \text{(ii)}$$

Applying the second law to loop acba gives

$$-12 + 7I_1 - 4I_3 = 0 \qquad \text{(iii)}$$

and to loop adba gives

$$-9 - 8I_2 - 4I_3 = 0 \qquad \text{(iv)}$$

We have three equations and the three unknowns I_1, I_2 and I_3.

From (iv) $\quad I_3 = -2I_2 - 2.25$

Substituting this in (iii) gives

$$-12 + 7I_1 + 9 + 8I_2 = 0; \quad \text{or,} \quad 7I_1 + 8I_2 = 3$$

Also substituting the value of I_3 in (ii) gives

$$I_1 - 2I_2 - 2.25 = I_2; \quad \text{or,} \quad I_1 = 3I_2 + 2.25$$

Substituting this value of I_1 in the previous equation,

$$21I_2 + 15.75 + 8I_2 = 3 \quad \text{or,} \quad I_2 = -0.44\text{A}.$$

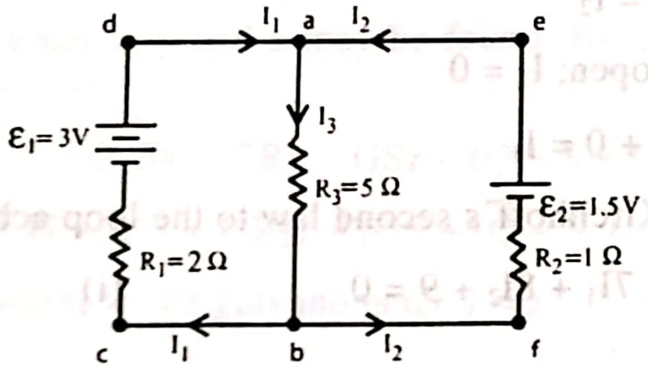
Using this value of I_2 in equation for I_1 gives

$$I_1 = 3(-0.44) + 2.25 = -1.32 + 2.25 = 0.93 \text{ A}.$$

Substituting the values of I_1 and I_2 in eqn. (ii), we get

$$I_3 = I_2 - I_1 = (-0.44) - 0.93 = -1.37\text{A}.$$

***Example 6.2** Calculate the currents flowing through the various branches of the circuit shown below.*



Soln.

Let us consider the closed circuit abcd. Applying Kirchhoff's second law, we get

$$-I_3R_3 - I_1R_1 + \varepsilon_1 = 0$$

or, $\varepsilon_1 = I_1R_1 + I_3R_3$

or, $3 = 2I_1 + 5I_3$ (i)

Again applying the second law to the closed circuit bcdaefb, we get

$$-I_1R_1 + \varepsilon_1 - \varepsilon_2 + I_2R_2 = 0$$

or, $\varepsilon_1 - \varepsilon_2 = I_1R_1 - I_1R_2$

or, $3 - 1.5 = 2I_1 - I_2$ (ii)

Applying the first law to the point b, we get

$$I_1 + I_2 - I_3 = 0 \quad \text{(iii)}$$

Eliminating I_3 with the help of eqn. (iii), we get from eqn. (i)

$$3 = 2I_1 + 5(I_1 + I_2)$$

or, $3 = 7I_1 + 5I_2$ (iv)

From eqn. (i) we have

$$1.5 = 2I_1 - I_2$$

or, $7.5 = 10I_1 - 5I_2$ (v)

Adding (iv) and (v) we get

$$10.5 = 17I_1$$

or, $I_1 = \frac{10.5}{17} = 0.617$ A.

Substituting the value of I_1 in eqn. (iv) we get

$$3 = 4.319 + 5I_2$$

or, $5I_2 = 3 - 4.319 = 0 - 1.319$

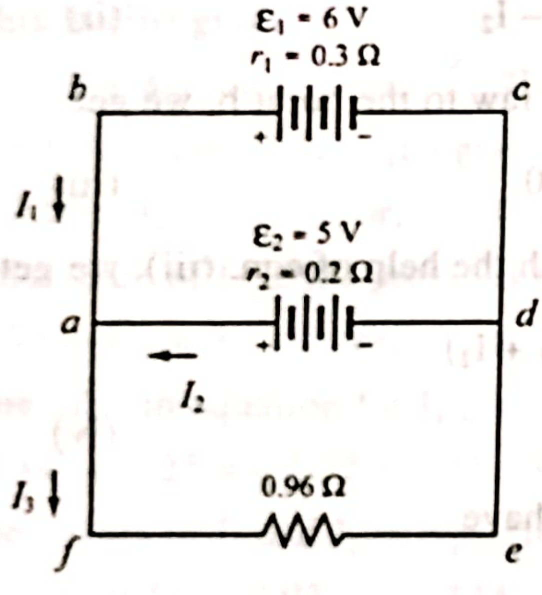
or, $I_2 = -\dfrac{1.319}{5} = -0.263A.$

From eqn. (iii) we get

$$I_3 = I_1 + I_2 = (0.617) + (-0.263)$$

$$= 0.354A.$$

Example 6.3 For the circuit shown in the figure below, find the current in the 0.96Ω resistor and the terminal voltages of the batteries.



Soln.

Applying Kirchhoff's first law at point a, we have

$$I_1 + I_2 - I_3 = 0 \qquad \text{(i)}$$

Applying Kirchhoff's second law to the loop abcda and moving counterclockwise, we get

$$-5 + (0.2)\,I_2 + 6 - (0.3)I_1 = 0$$

$$1 = 0.3I_1 - 0.2I_2$$

or, $3I_1 - 2I_2 = 10 \qquad$ (ii)

and to the loop afeda and again moving counterclockwise.

$$-(0.96)I_3 + 5 - (0.2)I_2 = 0$$

$$0.96\, I_3 + 0.2I_2 = 5$$

or, $9.6\, I_3 + 2I_2 = 50$ (iii)

Substituting (i) into (iii) to eliminate I_3 we get

$$9.6I_1 + 11.6I_2 = 50 \quad \text{(iv)}$$

Multiplying (ii) by 5.8, we get

$$17.4\, I_1 - 11.6\, I_2 = 58 \quad \text{(v)}$$

Adding (iv) and (v) we get

$$27I_1 = 108$$

or, $I_1 = 4A$.

Substituting the value of I_1 into (ii) we get

$$12 - 2I_2 = 10$$

or, $I_2 = \frac{2}{2} = 1.0A$.

Substituting the values of I_1 and I_2 into (i) we get

$$4 + 1 - I_3 = 0$$

or, $I_3 = 5A$.

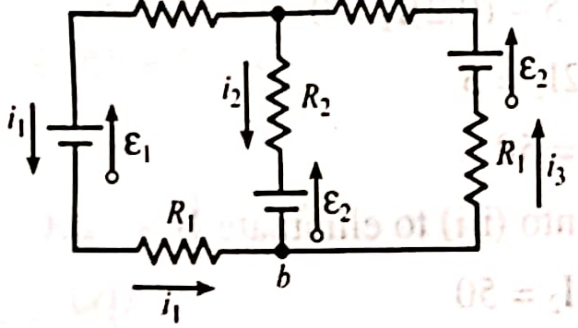
Terminal voltage V_1 is given by

$$V_1 = \varepsilon_1 - I_1r_1 = 6V - (4.0)\,(0.3)V$$
$$= 6V - 1.2V = 4.8V$$

and V_2 is given by

$$V_2 = \varepsilon_2 - I_2r_2$$
$$= 5V - (1.0)\,(0.2)\,V$$
$$= 5V - 0.2V$$
$$= 4.8V.$$

***Example 6.4** Find the currents in each of the three branches of the circuit given below and the potential difference between points a and b in the circuit.*



Given $\quad \varepsilon_1 = 2.1V \qquad \varepsilon_2 = 6.3V$

$\qquad\quad R_1 = 1.7\Omega \qquad R_2 = 3.5\Omega$

The directions of the currents have been chosen arbitrarily.

Soln.

Currents in each branches of the circuit

Applying the junction rule at a, we obtain

$$i_3 = i_1 + i_2 \qquad \text{(i)}$$

Starting from point a, let us traverse the left-hand loop in a counter clockwise direction. We obtain

$$-i_1R_1 - \varepsilon_1 - i_1R_1 + \varepsilon_2 + i_2R_2 = 0$$

$$\text{or,}\quad 2i_1R_2 - i_2R_2 = \varepsilon_2 - \varepsilon_1 \qquad \text{(ii)}$$

Traversing the right hand loop in a clockwise direction from point a, we find

$$+i_3R_1 - \varepsilon_2 + i_3R_1 + \varepsilon_2 + i_2R_2 = 0$$

$$\text{or,}\quad i_2R_2 + 2i_3R_1 = 0 \qquad \text{(iii)}$$

we have three independent simultaneous equations [(i), (ii) and (iii)] and three variables i_1, i_2 and i_3.

Substituting the value of i_3 from (i) into (iii)

$$i_2R_2 + 2(i_1 + i_2)R_1 = 0$$

$$\text{or,}\quad i_2R_2 + 2i_1R_1 + 2i_2R_1 = 0 \qquad \text{(iv)}$$

Subtracting (ii) from (iv)

$$2i_2(R_1 + R_2) = -(\varepsilon_2 - \varepsilon_1)$$

$$\text{or,}\quad i_2 = -\frac{\varepsilon_2 - \varepsilon_1}{2(R_1 + R_2)} = -\frac{(6.3V - 2.1V)}{2(1.7\Omega + 3.5\Omega)} = -0.40A.$$

Substituting the value of i_2 in (ii)

$$2i_1 (1.7\Omega) - (-0.40A)(3.5\Omega) = 6.3V - 2.1V$$

$$\text{or,}\quad 3.4\, i_1 + 1.4 = 4.2$$

$$\text{or,}\quad 3.4 i_1 = 4.2 - 1.4$$

$$\text{or,}\quad i_1 = \frac{2.8}{3.4} = 0.82A$$

and from (i) we get

$$i_3 = i_1 + i_2 = 0.82A + (-0.40A)$$

$$= 0.42A.$$

The signs of currents tell us that the directions of i_1 and i_3 have been guessed correctly but the direction of i_2 is wrong. It should point up not down.

For the potential difference between a and b, let us traverse the branch ab. We get

$$V_a - i_2R_2 - \varepsilon_2 = V_b$$

$$\text{or,}\quad V_a - V_b = \varepsilon_2 + i_2R_2$$

$$= 6.3\ V + (0 - 0.40A)(3.5\Omega)$$

$$= +4.9\ V.$$

6.4 RC circuit

2012-13

We have so for considered circuits containing only resistors, in which the currents did not vary in time. Now we shall introduce the capacitor as a circuit element, which will lead to the study of time–varying currents.

Let the capacitor C in Fig. 6.5 be charged by throwing the switch S to position a. As soon as the circuit is closed, charge would flow out from the positive terminal of the source of emf and be gradually accumulated in

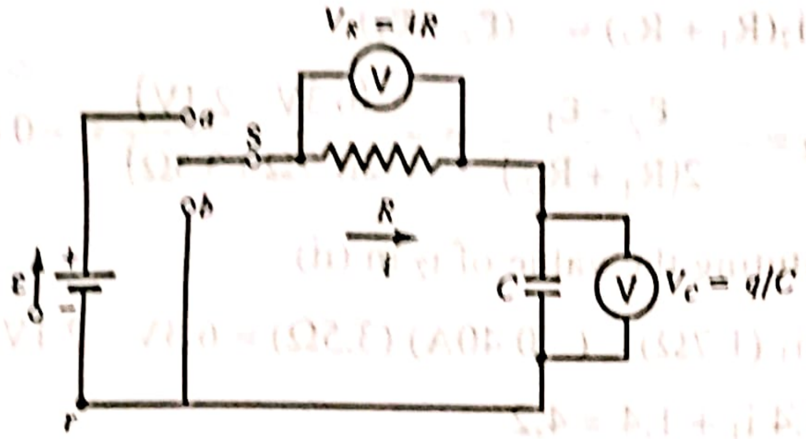


Fig. 6.5

the capacitor C. The current set up in the circuit can be determined by applying the principle of conservation of energy.

In a certain time dt an amount of charge dq (= idt) moves through any cross-section of the circuit. The work done by the seat of emf to move this charge is $\mathcal{E}dq$ where $\mathcal{E}$ is the emf of the battery. The internal energy produced in the resistor R during this time dt is i^2Rdt. At the same time an amount of energy dU will be stored in the capacitor where $U = \frac{q^2}{2C}$. Conservation of energy demands that

$$\mathcal{E}dq = i^2Rdt + dU$$

$$= i^2Rdt + d\left(\frac{q^2}{2C}\right)$$

$$\text{or,}\quad \mathcal{E}dq = i^2Rdt + \frac{2q}{2C}dq$$

$$= i^2Rdt + \frac{q}{C}dq$$

Dividing by dt, we obtain

$$\mathcal{E}\frac{dq}{dt} = i^2R + \frac{q}{C}\frac{dq}{dt}$$

Since q is the charge on the upper plate, positive i means positive $\frac{dq}{dt}$. With $i = \frac{dq}{dt}$, the above equation becomes

$$\mathcal{E} = iR + \frac{q}{C} \qquad \text{(i)}$$

[Eqn. (i) can also be obtained by applying loop theorem. Starting from point x and going around the circuit clockwise, we get

$$\mathcal{E} - iR - \frac{q}{C} = 0$$

or, $\mathcal{E} = iR + \frac{q}{C}$

same as eqn. (i)]

To solve eqn. (i) let us substitute $i = \frac{dq}{dt}$. Eqn. (i) then becomes

$$\mathcal{E} = R\frac{dq}{dt} + \frac{q}{C} \qquad \text{(ii)}$$

Rearranging, eqn. (ii) can be written as

$$R\frac{dq}{dt} = \mathcal{E} - \frac{q}{C}$$

or, $CR\frac{dq}{dt} = C\mathcal{E} - q$

or, $\frac{dq}{C\mathcal{E} - q} = \frac{dt}{CR}$ (iii)

Integrating eqn. (iii) we get

$$\int\frac{dq}{C\mathcal{E} - q} = \frac{1}{CR}\int dt$$

$$-\log_e (C\mathcal{E} - q) + A = \frac{t}{CR} \qquad \text{(iv)}$$

where A is a constant of integration.

At $t = 0$, $q = 0$

$\therefore \quad -\log_e C\mathcal{E} + A = 0$

or, $A = \log_e C\varepsilon$

Substituting this value of A in (iv), we get

$$-\log_e (C\varepsilon - q) - \log_e C\varepsilon = \frac{t}{CR}$$

$$\log_e (C\varepsilon - q) - \log_e C\varepsilon = -\frac{t}{CR}$$

or, $\log_e \dfrac{C\varepsilon - q}{C\varepsilon} = -\dfrac{t}{CR}$

or, $\dfrac{C\varepsilon - q}{C\varepsilon} = e^{-\frac{t}{CR}}$

or, $C\varepsilon - q = C\varepsilon e^{-\frac{t}{RC}}$

or, $q = C\varepsilon\left(1 - e^{-\frac{t}{RC}}\right)$ (v)

But $C\varepsilon = q_0$ (maximum charge)

$$\therefore \quad q = q_0\left(1 - e^{-\frac{t}{RC}}\right) \qquad (6.7)$$

Eqn. 6.7 gives the charge across the capacitor at any instant of time t.

The potential across the capacitor is given by

$$E = \frac{q}{C} = \frac{q_0}{C}\left(1 - e^{-\frac{t}{RC}}\right) = \varepsilon\left(1 - e^{-\frac{t}{RC}}\right) \qquad \text{(vi)}$$

and the current by

$$i = \frac{dq}{dt} = \frac{d}{dt}\left[q_0\left(1 - e^{-\frac{t}{RC}}\right)\right]$$

$$= \frac{q_0}{RC}e^{-\frac{t}{RC}} = \frac{\varepsilon}{R}e^{-\frac{t}{RC}} = i_0 e^{-\frac{t}{RC}} \qquad (6.8)$$

where $i_0 = \dfrac{\varepsilon}{R}$ is the maximum current.

As can be seen from eqn. (6.7) that the charge on a capacitor grows exponentially to its final value q_0. At $t = 0$, $V_R = \varepsilon$ and $V_C = 0$ *i.e.*, the capacitor is not charged and the full potential difference appears across R. As $t \rightarrow \infty$, $V_C \rightarrow \varepsilon$ (the capacitor becomes fully charged) and $V_R \rightarrow 0$; the current stops flowing. At all times $\varepsilon = V_R + V_C$ as eqn. (ii) requires.

The quantity RC in eqns. (6.7), (6.8) and (6.9) has the dimensions of time (the exponent must be dimensionless) and is called the *capacitive time constant* τ_C of the circuit.

$$\tau_C = RC \qquad \text{(ix)}$$

$$\left[RC = \frac{V}{i}C = \frac{V}{i}\cdot\frac{q}{V} = \frac{V}{q/t}\cdot\frac{q}{V} = t\right]$$

If t is numerically equal to RC, then eqn. (6.7) becomes

$$q = q_0\,(1 - e^{-1}) = q_0\left(1 - \frac{1}{2.72}\right) = 0.638q_0$$

$$= 0.638C\varepsilon.$$

Hence in RC seconds, the charge on the condenser reaches 63% of its final value q_0. *The time constant of the circuit is therefore defined as the time taken by a condenser to reach 63 per cent of its final charge.*

A condenser takes time to be fully charged to its fullest extent. The growth of charge in a condenser with time is shown in Fig. 6.6.

The curve in the figure is an exponential curve. As can be seen, from the figure above, the increase of the charge on the capacitor towards its limiting value is delayed by a time characterized by the time constant RC. With no resistor present (RC = O), the charge would rise immediately to its limiting value. When the switch is closed on "a", the charge on the capacitor is initially zero, so the potential difference (p.d) across the capacitor is initially zero. At this time $\varepsilon = iR$ (eqn. ii) and so $i = \varepsilon/R$ at $t = 0$. Because of the current charge flows to the capacitor and the p.d. across the capacitor

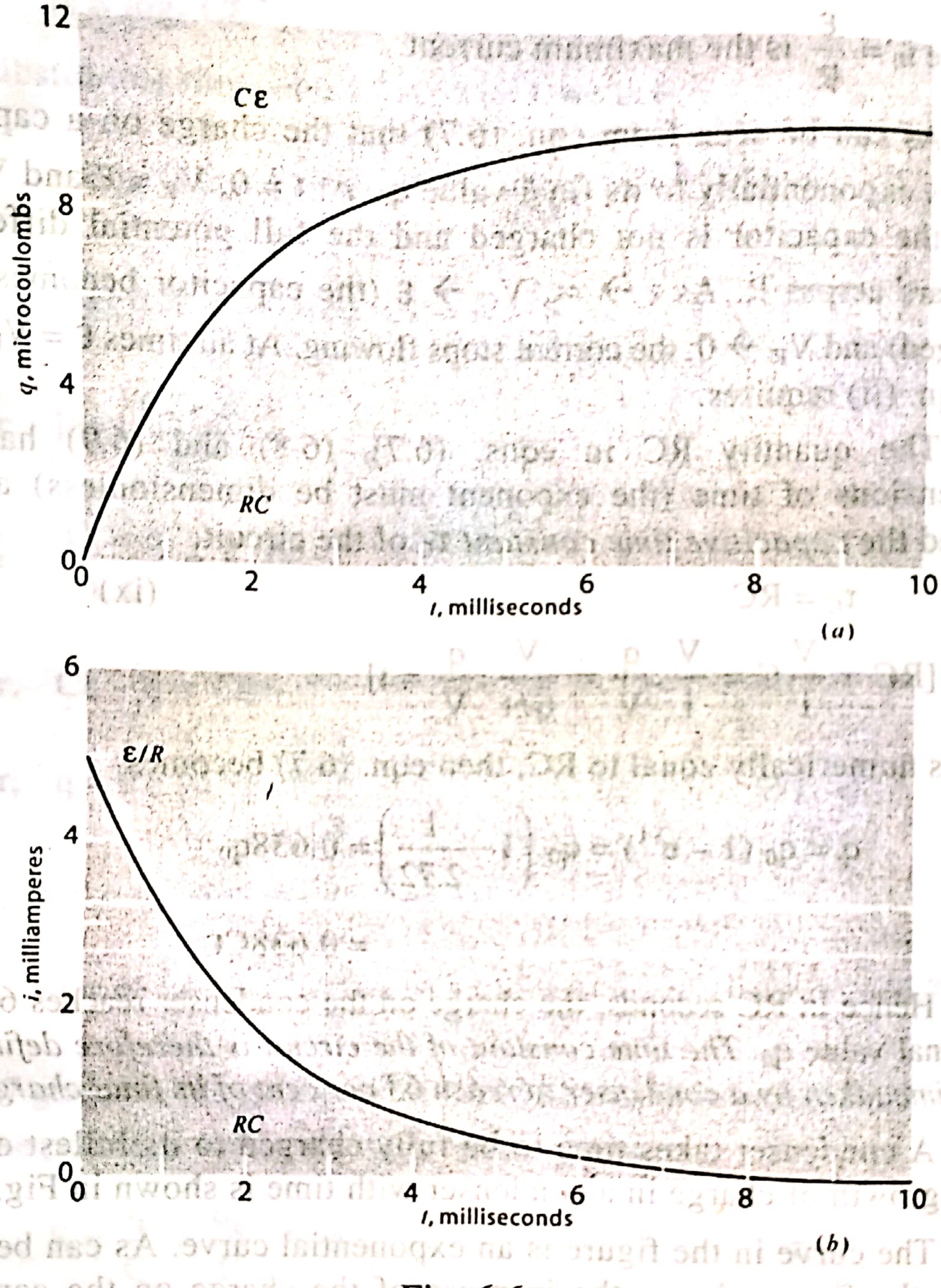


Fig. 6.6

increases with time. As can be seen from eqn. (ii), because the emf ℰ is constant, any increase in p.d. across the capacitor must be balanced by a corresponding decrease in the p.d across the resistor with a similar decrease in the current. This decrease in current means that the charge on the capacitor increases more slowly. This process continues until the current decreases to zero at which time there is no p.d across the resistor. The entire p.d. of the emf now appears across the capacitor, which is fully charged

($q = C\mathcal{E}$). Unless changes are made in the circuit there is no further flow of charge.

Discharging of a capacitor

Assume that the switch S (Fig. 6.5) has been in position 'a' for a time that is much greater than RC. For all practical purpose the capacitor is fully charged, and no charge is flowing. Now throw the switch S to position 'b'. How do the charge of the capacitor and the current vary with time?

As the switch S is closed on b, the capacitor discharges through the resistance R – the charge now flows out of the capacitor. The current now flows in the opposite direction and hence the potential difference across the resistor is negative compared with its value during the charging process. There is no emf in the circuit, *i.e.*, $\mathcal{E} = 0$. Eqn. (i) therefore reduces to

$$iR + \frac{q}{C} = 0$$

$$\text{or,}\quad \frac{q}{C} = -iR = -\frac{dq}{dt}R$$

$$\text{or,}\quad \frac{dq}{q} = -\frac{dt}{RC}$$

Integrating

$$\log_e q = -\frac{t}{RC} + K$$

where K is a constant of integration.

This time

$$q = q_0 \quad \text{at} \quad t = 0$$

Putting these values , we get

$$\log_e q_0 = K$$

$$\text{or,}\quad \log_e q = -\frac{t}{RC} + \log_e q_0$$

$$\text{or,} \quad \log_e \frac{q}{q_0} = -\frac{t}{RC}$$

$$\text{or,} \quad \frac{q}{q_0} = e^{-\frac{t}{RC}}$$

$$\text{or,} \quad q = q_0 e^{-\frac{t}{RC}} \tag{6.10}$$

Eqn. (6.10) gives the charge on the capacitor at any time t during discharge.

Dividing by C, we get

$$\frac{q}{C} = \frac{q_0}{C} e^{-t/RC}$$

$$\text{or,} \quad E = \varepsilon e^{-t/RC} \tag{6.11}$$

Eqn. (6.11) gives the expression for p.d. across C during discharge.

Again

$$\frac{dq}{dt} = \frac{d}{dt}\left(q_0 e^{-\frac{t}{RC}}\right)$$

$$= -\frac{q_0}{RC} e^{-\frac{t}{RC}}$$

$$\text{or,} \quad i = -\frac{\varepsilon}{R} e^{-\frac{t}{RC}} = -i_0 e^{-\frac{t}{RC}} \tag{6.12}$$

The negative sign shows that the current is in the opposite direction than that during charging. This is as it should be since the capacitor is discharging rather than charging. Since $q_0 = C\varepsilon$, we can write

$$i = -\frac{\varepsilon}{R} e^{-\frac{t}{RC}}$$

The initial current found by setting $t = 0$ is $-\frac{\varepsilon}{R}$. This is reasonable because the initial p.d. across the resistor is ε.

The p.d across R and C, which are respectively proportional to i and q, can again be measured as indicated in the circuit diagram. Typical results are shown in Fig. 6.7 above. It can be noted that as V_c (q/C) falls exponentially from its maximum value, which occurs at $t = 0$, V_R (= iR) which is negative, rises exponentially to zero. As can be seen $V_c + V_R = 0$ as recquired by loop theorem.

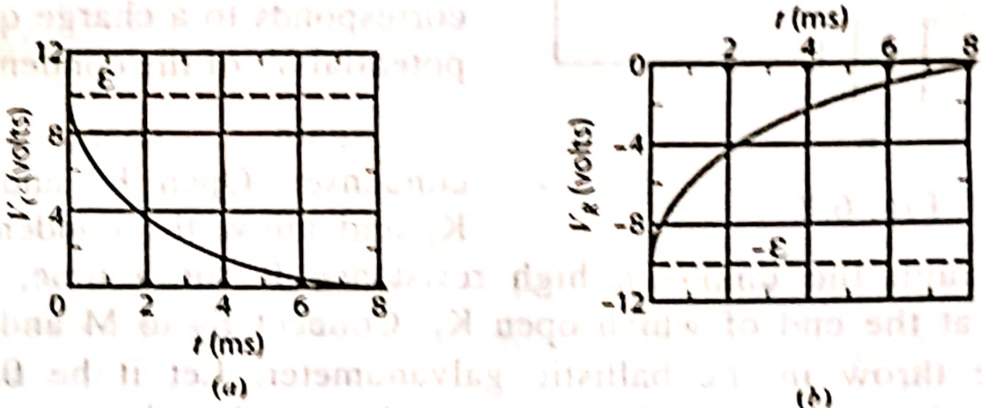


Fig. 6.7

Substituting t = RC in eqn. (6.10), we get

$$q = q_0 e^{-1} = q_0 \frac{1}{2.72} = 0.367 q_0$$

Thus the time constant during discharge is defined as the time $t = \tau_c = RC$ for the capacitor charge to be reduced to 37 per cent of the initial charge q_0.

6.5 Application of discharge of a condenser

(i) Measurement of high resistance

The rate of discharge of a condenser can be used to measure high resistances, of the order of 10^6 ohms or more, *e.g.*, the insulation resistance of a cable or that of a chiana insulator. A circuit similar to one shown in Fig. 6.8 is usually used.

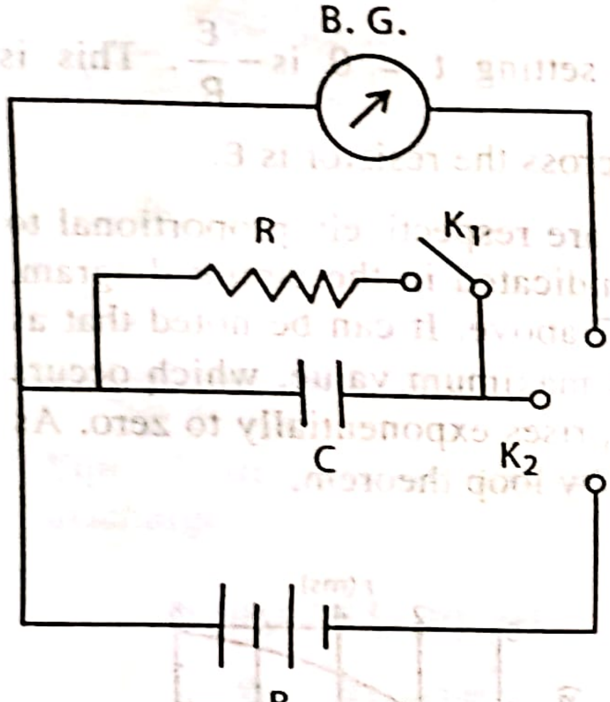


Fig. 6.8

B.G. is a ballistic galvanometer, R is the unknown high resistance, C is a condenser of known capacity and B is a battery.

Keep K_1 open. Press K_2 to N when the condenser is charged to a potential V_0. Next connect K_2 to M and note the throw in the ballistic galvanometer. Let it be θ_0. This corresponds to a charge q_0 for a potential V_0 of the condenser.

Again charge the condenser. Open K_2 and close K_1 and allow the condenser to leak through the unknown high resistance R for a time, say, t seconds at the end of which open K_1. Connect K_2 to M and again note the throw in the ballistic galvanometer. Let it be θ. This corresponds to a charge q left in the condenser when the potential V_0 of the condenser has dropped to a value V after t seconds due to the discharge.

From the relation for discharge of condenser (eqn. 6.10), we have

$$q = q_0 e^{-\frac{t}{RC}}$$

$$\text{or,} \quad \frac{q}{q_0} = e^{-\frac{t}{RC}}; \quad \text{or,} \quad \frac{q_0}{q} = e^{\frac{t}{RC}}$$

$$\text{or,} \quad \frac{t}{RC} = \log_e \frac{q_0}{q}; \quad \text{or,} \quad RC = \frac{t}{\log_e \frac{q_0}{q}}$$

$$\text{or,} \quad R = \frac{t}{C \log_e \frac{q_0}{q}}$$

$$\text{But } \frac{q_0}{q} = \frac{V_0}{V} = \frac{\theta_0}{\theta}$$

$$R = \frac{t}{C \log_e \dfrac{\theta_0}{\theta}}$$

If t is in seconds, C is in farads, R is in ohms. If C is known R can be calculated or vice versa. A series of observations for t an θ are obtained and a graph between t and $\log_e \dfrac{\theta_0}{\theta}$ is drawn which should be a straight line. From this the value of R may be obtained. Instead of using a ballistic galvanometer to measure the charge, a suitable device (a digital voltmeter) may be used for recording the voltage directly. For example a charged capacitor of known capacitance C may be connected across the unknown resistance and allowed to discharge. The lime t at which the potential difference across the capacitor drops to 0.37 of its original value is noted. Then RC = t seconds from which the value of the unknown resistance may be obtained.

(ii) Leakage resistance of a dielectric

Due to imperfection of the dielectric of the condenser used, there may be natural leakage of the condenser. To test this, the condenser is charged and then immediately discharged through the ballistic galvanometer and the throw is noted. It is then recharged momentarily and after lapse of about 1 to 10 minutes, allowed to discharge through the galvanometer. If the throw in the galvanometer is different form that obtained before, there is a leakage in the dielectric of the condenser.

To determine the leakage resistance of the dielectric, the condenser is charged by connecting the switch K_2 to N (Fig. 6.8). K_2 is then immediately connected to M and the throw θ_0 in the ballistic galvanometer is noted. K_2 is then again connected to N, C is then isolated from both N and M for a measured time t; and finally K_2 is connected to M. Let the throw be θ_1. The insulation resistance R_1 is given by

$$R_1 = \frac{t}{C\log_e \frac{\theta_0}{\theta_1}}$$

To find the unknown resistance R when there is leakage in the condenser, connect K_2 to N, then isolate C from N and M for a definite time t_1 during which interval K_1 is closed to allow the condenser to leak across R. K_1 is then opened and K_2 connected to M. Let θ_2 be the throw in the galvanometer. Under these circumstances, the resistance across the condenser is

$$\frac{1}{R} + \frac{1}{R_1} = \frac{RR_1}{R+R_1}$$

$$\frac{RR_1}{R+R_1} = \frac{t_2}{C\log_e \frac{\theta_0}{\theta_2}}$$

Thus R can be determined.

Example 6.5 After how many time constants will the energy stored in a capacitor reach one-half its equilibrium value?

Soln.

The energy of a charged capacitor is given by

$$U = \frac{1}{2}\frac{q^2}{C}$$

the equilibrium energy U_∞ being $\frac{1}{2C}(C\varepsilon)^2$ where ε is the emf to which the capacitor has been charged.

$$U = \frac{1}{2C}(C\varepsilon)^2\left(1-e^{-\frac{t}{RC}}\right)^2$$

$$U = U_\infty\left(1-e^{-\frac{t}{RC}}\right)^2$$

Putting $U = \frac{1}{2} U_\infty$, we get

$$\frac{1}{2} = \left(1 - e^{-\frac{t}{RC}}\right)^2$$

or, $0.707 = (1 - e^{-t/RC})$

$$e^{-\frac{t}{RC}} = 0.293; \qquad e^{\frac{t}{RC}} = 3.413$$

$$\frac{t}{RC} = \ln(3.413) = 1.28 \quad \text{or,} \quad t = 1.28\,(RC).$$

Example 6.6 A 150 capacitor is connected through a 500–Ω resistor to 40–V battery. (a) what is the time constant of the circuit? (b) what is the final charge q_0 on a capacitor plate? (c) How long does it take for the charge on a capacitor plate to reach $0.8q_0$?

Soln.

(a) $t = RC = (500\Omega)(150 \times 10^{-6}F)$

$= 75ms.$

(b) $q_0 = VC = (40V)(150 \times 10^{-6}F)$

$= 6.0 \times 10^{-3}C$

$= 6.0$ mC.

(c) $q = q_0 (1 - e^{-\frac{t}{RC}})$

$0.8 = 1 - e^{-t/RC}$

$$e^{\frac{t}{RC}} = \frac{1}{0.2} = 5; \qquad \frac{t}{RC} = \ln 5 = 1.6$$

$\therefore \quad t = (RC)(1.6) = (75)(1.6)$ ms

$= 120$ ms.

Example 6.7 *A 50μF capacitor initially uncharged is connected through a 300–Ω resistor to a 12–V battery. (a) What is the magnitude of the final charge q_0 on the capacitor? (b) What is the time constant of the circuit? (c) How long after the capacitor is connected to the battery will it be charged to $\frac{1}{2}q_0$ and to 0.90 q_0?*

Soln.

(a) $q_0 = CV = (50 \times 10^{-6}\,F)(12V)$

$= 600 \times 10^{-6}\,C = 600\,\mu C.$

(b) $\tau = RC = (300\,\Omega)(50 \times 10^{-6}F)$

$= 15 \times 10^{-3}s = 15ms.$

(c) $q = q_0\left(1 - e^{-\frac{t}{RC}}\right)$

or, $\frac{q_0}{2} = q_0\left(1 - e^{-\frac{t}{RC}}\right)$

or, $\frac{1}{2} = 1 - e^{-\frac{t}{RC}}$

or, $e^{-\frac{t}{RC}} = \frac{1}{2}$

or, $e^{+\frac{t}{RC}} = 2$

$\frac{t}{RC} = \ln 2 = 0.693$

$\therefore\ t = (0.693)(RC) = (0.693)(15\text{ ms})$

$= 10.5\text{ ms}.$

Similarly

$0.9\,q_0 = q_0\left(1 - e^{-\frac{t}{RC}}\right)$

or, $e^{-\frac{t}{RC}} = 0.1$; or, $e^{-\frac{t}{RC}} = 10$

$$\frac{t}{RC} = \ln 10 = 2.30$$

$$\therefore \quad t = (2.30)\,(RC) = (2.30)\,(15\text{ ms}) = 34.5\text{ ms.}$$

Example 6.8 In a laboratory, a student charges a 2-μF capacitor by placing it across a 1.5-V battery. While disconnecting it, the student holds its two lead wires in two hands. Assuming that the resistance of the body between the two hands is 60kΩ, what is the time constant of the series circuit composed of the capacitor and the student's body? How long does it take for the charge to drop to $1/e$ *of its original value and to* $\frac{1}{100}$*?*

Soln.

$$t = RC = (60 \times 10^3\ \Omega)\,(2 \times 10^{-6}\ F) = 120 \times 10^{-3}\ s = 0.12s.$$

$$q = q_0 e^{-\frac{t}{RC}}; \qquad \frac{q_0}{e} = q_0 e^{-\frac{t}{RC}}$$

$$e^{-\frac{t}{RC}} = \frac{1}{e}; \qquad e^{\frac{t}{RC}} = e$$

$$\frac{t}{RC} = \ln(e) = \ln 2.728 = 1.00$$

or, $t = (1.00)\,(RC) = (1.00)\,(0.12s) = 0.12s.$

Similarly,

$$\frac{q_0}{100} = q_0\, e^{-\frac{t}{RC}}$$

or, $e^{-\frac{t}{RC}} = \frac{1}{100}; \qquad e^{\frac{t}{RC}} = 100$

$$\frac{t}{RC} = \ln(100) = 4.605$$

$$\therefore \quad t = (RC)(4.605) = (0.12s)(4.605)$$

$$= 0.55 \text{ s.}$$

Example 6.9 A 1-μF capacitor is allowed to discharge through an unknown resistance. If the charge on the capacitor takes 34.65 seconds to drop to half of its original value, what is the value of the resistance?

Soln.

$$q = q_0 e^{-\frac{t}{RC}}$$

$$e^{-\frac{t}{RC}} = \frac{q}{q_0} = \frac{\frac{1}{2}q_0}{q_0} = \frac{1}{2}$$

$$e^{\frac{t}{RC}} = 2; \qquad \text{or,} \quad \frac{t}{RC} = \ln.2 = 0.693$$

$$\text{or,} \quad (0.693)\,RC = t = 34..65$$

$$RC = \frac{34.65}{0.693} = 50$$

$$R = \frac{50}{C} = \frac{50}{10^{-6}} = 50 \times 10^{6} \text{ ohms.}$$

EXERCISES

1. State and explain Kirchhoff's laws for the distribution of currents in network of conductors.

2. State Kirchhoff's laws relating to the distribution of current in a network of conductors. Apply the laws to a Wheatstone bridge and find the conditional for bridge balance.

3. Discuss in detail the statement that the energy method and the loop rule method for solving circuits are perfectly equivalent.

4. Obtain an expression for the growth of charge when a capacitor is charged through a resistance for a constant emf. What is time constant of the circuit?

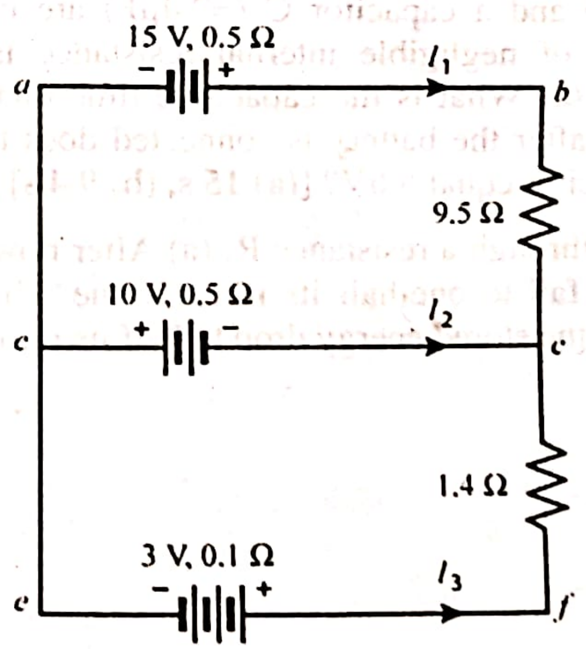
5. Find the expression for current i in a RC circuit for a constant emf E at any instant after making the circuit.

6. A capacitor is charged for a constant emf. and then allowed to discharge through a resistance. Find the expression for the discharge of the condenser.

7. What is time constant of a RC circuit? Show that the product RC has the dimension of time, *i.e.*, 1 second = 1 ohm × 1 farad.

8. Describe the method of leakage for measuring high resistance. Deduce any formula used.

9. Find I_1, I_2, I_3 and the potential difference from point b to point e in the figure shown below:



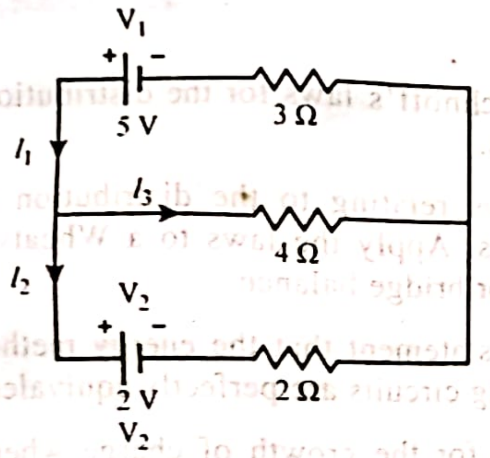
$I_1 = 2A$

$I_2 = 8A$

$I_3 = 6A$

$V_e - V_b = -13V$

10. Find the current in each branch of the circuit shown below:



$I_1 = \frac{11}{13}$ A

$I_2 = \frac{3}{13}$ A

$I_3 = \frac{8}{13}$ A

11. A 400 – μF capacitor is connected through a resistor to a battery. Find (a) the resistance R and (b) the emf ε of the battery if the time constant of the circuit is 0.5s and the maximum charge on the capacitor is 0.024C.

[R = 1250 Ω; ε = 60V]

12. In a certain electronic device, a 10 – μF capacitor is charged to 200V. When the device is shut off, the capacitor is discharged for safety reasons by a so – called bleeder resistor of 1MΩ placed across its terminals. How long does it take for the charge on the capacitor to decrease to 0.01 of its original value? **[46 s]**

13. A condenser of capacity 1 μF in discharged through a high resistance, the time taken for half the charge on the condenser to leak was found to be 10 sec. Compute the value of the resistance. **[15.8×10^6 ohms]**

14. A resistor R (= 6.2 MΩ) and a capacitor C (=2.4μF) are connected in series, and 12–V battery of negligible internal resistance is connected across their combination. (a) What is the capacitive time constant of the circuit? (b) At what time after the battery is connected does the potential difference across the capacitor equal 5.6V? **[(a) 15 s, (b) 9.4 s]**

15. A capacitor C discharges through a resistance R. (a) After how many time constants does its charge fall to one-half its initial value? (b) After how many time constants does the stored energy drop to half its initial value?

[0.69 τ, 0.35 τ]

CHAPTER - VII

THE MAGNETIC FIELD

In the preceding chapters we have studied the forces between electrostatic charges and other related phenomena. Nothing has been said about the magnetic forces, that is forces between magnetic bodies. Such forces have been known ever since the discovery of naturally occuring magnetite (Fe_3O_4) by the ancient Greeks. The science of magnetism grew from the observation that certain 'stones' like magnetite would attract bits of iron. The word *magnetism* comes form the district of Magnesia in Asia Minor, which is one of the places at which the stones were found. The forces were described in a manner similar to the treatment of electrostatic forces. Instead of electrostatic charges, the concept of magnetic poles were introduced. The poles were supposed to be of two types - *north* and *south* - which attract of repel each other according to Coulomb's law. All the laws of electrostatics can then be reproduced with magnetic poles replacing the electrostatic charges. But magnetic poles simply turned out to be mathematical constructs, they do not really exist.

It was Danish physicist Hans Christian Oersted who, in 1820, first linked the then separate sciences of current electricity and magnetism by showing that an electric current in a wire could deflect a magnetic compass needle. Subsequent works by Gauss, Henry, Faraday, Maxwell and others, have shown that magnetic and electric fields are inextricably interwined, that magnetic poles do not exist, and that the magnetic forces are nothing but forces between moving electric charges.

7.1 The Magnetic Field

The forces between two moving charges can be separated into two parts : one moving charge is said to create a *magnetic field* and this magnetic field then produces a magnetic force on the second moving charge, This magnetic field is more specifically called the field *of magnetic induction* and is usually denoted by **B**. This is similar to the electric field **E**. Now the electric field at a point could be measured by placing a test-charge q at that point and measuring

the electric force $\mathbf{F_E}$ that acts on the change. **E** is then defined by the relation

$$\mathbf{F_E} = q\,\mathbf{E} \quad (7.1)$$

If a magnetic monopole were available, **B** could have been defined in a similar way. Because such particles have not been found in nature, **B** must be defined in another way, in terms of the magnetic force exerted on a moving electric charge.

In principle, this in done by firing a test charge through a point where **B** is to be defined, using various speeds and directions for the test charge and determining the force (if any) that acts on the charge at that point. After many such trials it was found that the field of magnetic induction **B** is defined as the vector that satisfies the relation

$$\mathbf{F_B} = q\,\mathbf{v} \times \mathbf{B} \quad (7.2)$$

for all velocities **v** of the charge q which may be either positive or negative. In eqn. (7.2) some sort of limiting process is to be carried out so that the test charge q does not affect the sources of **B**. On analyzing eqn. (7.2) the following points emerge :

(i) The magnetic force $\mathbf{F_B}$ is not in the direction of the field, as in the case of electrostatics, but at right angles both to the field and to the velocity of the charge *i.e.*, magnetic force will always be a sideways fore. This means that a constant and a uniform magnetic force can neither speed up not slow down a moving charged particle; but can only deflect its direction of motion, that is the force can change only the direction of the particle's velocity **v** but not its magnitude. This may seem to violate Newton's second law of motion. But the law F = ma, concerns only vectors and when there in a change in velocity vector **v**, if only in direction, there is an acceleration. Since the magnitude of **v** does not change, the magnetic force does not change the kinetic energy of the particle. The actual direction of $\mathbf{F_B}$ can be determined by applying the right handed screw rule for determining the direction of the resultant vector obtained by vector (or cross) product of two vector. Fig. 7.1 shows the directions of the three vectors that appear in eqn. 7.2. It can be seen that $\mathbf{F_B}$, being at right angles to the plane formed by **v** and **B**, will always be at right angles to **v** (and also to **B**) and thus will always be a sideways force.

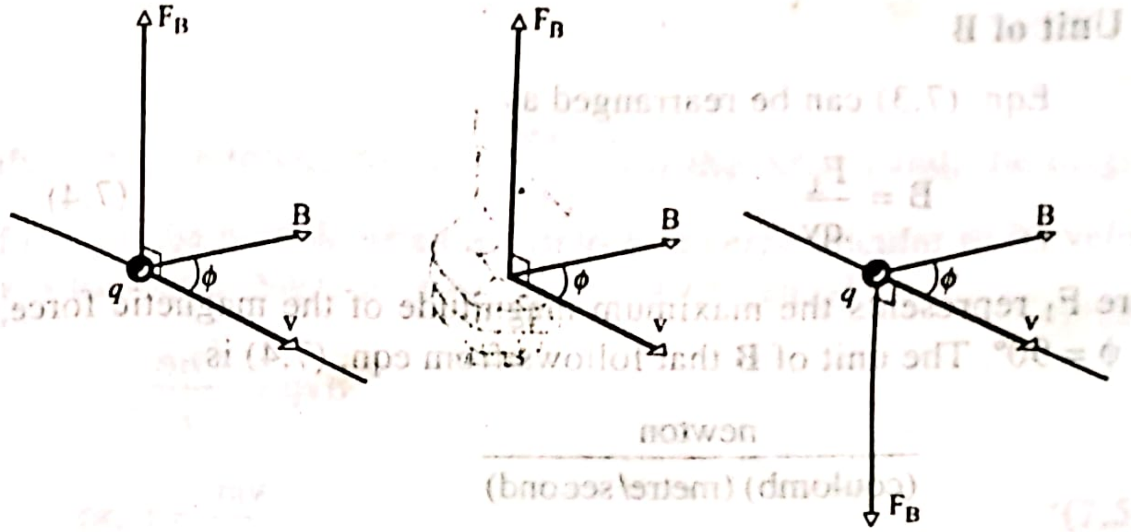


Fig. 7.1

The right hand rule in which **v** is *swept* into **B**, the outstretched thumb gives the direction of $\mathbf{F_B}$ for a positively charged particle. If the particle is negatively charged, the direction of $\mathbf{F_B}$ is *opposite* the thumb.

(ii) The magnitude of the deflecting magnetic force is given by

$$F_B = qvB \sin \phi \quad (7.3)$$

where ϕ is the angle between the directions of **v** and **B** and q is the absolute value of the charge. From this it can be seen that

(a) when **v** is either parallel or anti-parallel to **B**, which corresponds to $\phi = 0°$ or $180°$, F_B is indeed zero. The *maximum* value of the deflecting force (= qvB) occurs when the test charge is moving perpendicular to the magnetic field, *i.e.*, when $\phi = 90^0$.

(b) The magnetic force is directly proportional to q and v. The greater the charge of a particle and the faster it is moving, the greater the magnetic deflecting force. If the particle is stationary (v = 0) or is electrically neutral (q = 0), there is no such force acting on it.

(iii) The direction of the magnetic deflecting force depends on the nature of the charge q. If the charge is positive, the direction of the force is determined by using the right-hand rule for cross products. If the charge is negative, the direction is always opposite to that for the positive charge.

Unit of B

Eqn. (7.3) can be rearranged as

$$B = \frac{F_{\perp}}{qv} \tag{7.4}$$

where $F_{\perp}$ represents the maximum magnitude of the magnetic force, *i.e.*, $\phi = 90°$. The unit of **B** that follows from eqn. (7.4) is

$$\frac{\text{newton}}{(\text{coulomb})\,(\text{metre/second})}$$

This is given the special name *Weber/metre*² or *tesla*. Recalling that a coulomb per second is an ampere, we have.

$$1 \text{ tesla} = 1\text{ T} = 1\frac{\text{newton}}{(\text{coulomb/second})\,(\text{metre})}$$

$$= 1\frac{\text{N}}{\text{A.m}}$$

An earlier unit of B is the Gauss, its relationship with Weber/meter² (or tesla) being

$$1 \text{ Weber/metre}^2 = 1 \text{ tesla} = 10^4 \text{ Gauss.}$$

7.2 Magnetic force on a charge

Since the magnetic force $\mathbf{F_B}$ is always at right angles to the direction of motion, the work done by this force on the particle is zero. If the particle moves through an element of path dl, then the work done is $dW = \mathbf{F_B}.dl = 0$. Thus the magnetic field cannot change the kinetic energy of the particle, it can only deflect the particle sideways. If the field **B** is uniform and perpendicular to the velocity **v** then the magnitude of the sideways force is constant. The acceleration which it produces on the particle is also constant in magnitude and perpendicular to **v**. This is the characteristic of uniform circular motion. The path of the particle is, therefore, a circle and the magnetic force is a centripetal force.

If r be the radius of the circle and m be the mass of the particle,

then the centripetal force is $\frac{mv^2}{r}$. On the other hand, the magnetic force on the particle when the field **B** is perpendicular to its velocity v is $F_B = qvB$. Since the two forces must be equal

$$\frac{mv^2}{r} = qvB$$

$$\text{or,} \quad r = \frac{mv}{qB} \tag{7.5}$$

Thus the radius of the circle is directly proportional to the momentum mv of the particle and inversely proportional to the field B. The time taken by the particle to complete one circular revolution is

$$T = \frac{2\pi r}{v} \tag{7.6}$$

Substituting the value of r in (ii), we have

$$T = \frac{2\pi}{v} \cdot \frac{mv}{qB} = \frac{2\pi m}{qB} \tag{7.7}$$

Thus T is independent of r and v. Fast particle move in large circles (eqn. 7.5) and slower ones in small circles, but require the same time T (the *period*) to complete one revolution in the field.

The frequency f is

$$f = \frac{1}{T} = \frac{qB}{2\pi m}$$

The angular frequency w of the motion is

$$w = 2\pi f = \frac{qB}{m}$$

If the velocity v of the particle makes an angle θ with the field **B**, then the particle moves in a helix advancing with the velocity vcosθ in the direction of the field. The radius of a turn of the helix is equal to

$$r = \frac{mv\sin\theta}{qB} \tag{7.8}$$

The parallel component determines the pitch of the *helix* which is defined as the x-distance travelled during one revolution. Or

$$\text{pitch} = v_{11} \times T$$

$$= v_{11} \times \frac{2\pi r}{v_\perp}$$

$$= v \cos\theta \times \frac{2\pi r}{v \sin\theta} \tag{7.9}$$

It the charged particle enters into region in which both electric and magnetic fields are present, then the total force **F** acting on the particle is found by combining eqns. 7.1 and 7.2.

$$\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} \tag{7.10}$$

which is known as Lorentz force.

7.3 Magnetic lines of induction

A space around a magnet or a current carrying conductor can be regarded as the site of a magnetic field just as the space near a charged rod can be regarded as the site of an electric field. The basic magnetic field vector **B**, which has been defined in Art 7.1, can be represented by *lines of induction*, just as the electric field was represented by lines of force. The magnetic field vector is related to the lines of induction in the same manner as the electric field is related to the electric lines of force. That is

(i) The tangent to a line of induction, also known as a magnetic field line, at any point gives the direction of **B** at that point.

(ii) The lines of induction are drawn so that the number of lines of induction per unit cross-sectional area can be regarded as the measure of the magnitude of **B**. Where the lines are close together B is large and where they are far apart B is small. Like the electric field the field vector **B** is of fundamental importance. The lines of induction give a graphic representation of the way **B** varies throughout a certain region of space.

The flux ϕ_B for the magnetic field can be defined in exactly the same way as the flux ϕ_E for the electric field, namely

$$\phi_B = \int \mathbf{B}.d\mathbf{S}$$

in which the integral is taken over the surface (closed or open) for which ϕ_B is defined.

Example 7.1 An electron is moving with a speed of 5×10^7 m/s at right angle to a magnetic field of 5000 G. (a) What is the magnetic force on the electron? (b) What is the radius of the circle in which the electron moves?

Soln.

B = 5000 G = 0.5T

$q = 1.6 \times 10^{-19}$ C

$v = 5 \times 10^7$ m/s

(a) F = qvB

$$F = (1.6 \times 10^{-19}C)\ (5 \times 10^7\ m/s)\ (0.5T)$$

$$= 4.0 \times 10^{-12}\ N.$$

$m = 9.1 \times 10^{-31}$ kg

$v = 5 \times 10^7$ m/s

$F = 4.0 \times 10^{-12}$ N

(b) $F = \frac{mv^2}{r}$; or $r = \frac{mv^2}{F}$

$$r = \frac{(9.1\times10^{-31}\,kg)\,(5\times10^7\,m/s)^2}{4.0\times10^{-12}\,N}$$

$$= 5.7 \times 10^{-4} m = 0.57\ mm.$$

Example 7.2 Electrons enter a uniform magnetic field at right angles to the field (a) What is the magnitude of the field for electrons which traverse a complete circle in 10^{-9} sec? (b) What is the radius of curvature of the circle if each electron has an energy of 2000 eV?

Soln.

(a) From $T = \frac{2\pi m}{qB}$, we have

$$B = \frac{2\pi m}{qT} = \frac{(2)\,(3.141)\,(9.1\times10^{-31}\,kg)}{(1.6\times10^{-19}\,C)\,(10^{-9}\,sec)}$$

$$= 0.036\ W/m^2 = 0.036\ T.$$

(b) $2000\ eV = (2000)\ (1.6 \times 10^{-19}\ J)$

$$= 3.20 \times 10^{-16}\ J.$$

$$\frac{1}{2}mv^2 = K.E. = 3.20 \times 10^{-16}\ J$$

$$v^2 = \frac{2\times3.20\times10^{-16}\ J}{(9.1\times10^{-31}\ kg)} = 7.03 \times 10^{14}\ m^2/s^2$$

$$v = 2.65 \times 10^7\ m/s$$

$$\therefore\ r = \frac{mv}{qB} = \frac{(9.1\times10^{-31}\ kg)(2.65\times10^7\ m/s)}{(1.6\times10^{-19}\ C)(0.036\ W/m^2)}$$

$$= 4.2\ mm.$$

***Example 7.3** (a) What is the force on a singly charged carbon ion moving with a speed of 3×10^5 m/s at right angles to a magnetic field of 7500G ? (b) What is the centripetal acceleration of the ion? What is the radius of the circle in which the ion moves ?*

Soln.

(a) $q = e = 1.6 \times 10^{-19}\ C \quad B = 7500\ G = 0.75\ T$

$$F = Bqv = (0.75\ T)(1.6 \times 10^{-19} C)(3 \times 10^5\ m/s)$$

$$= 3.6 \times 10^{-14} N.$$

(b) $m_C = (12\ u)(1.66 \times 10^{-27} kg/u) = 19.9 \times 10^{-27}\ kg$

$$a = \frac{F}{m_C} = \frac{3.6\times10^{-14}\ N}{19.9\times10^{-27}\ kg} = 1.81 \times 10^{12}\ m/s^2.$$

(c) $$F = \frac{mv^2}{r}$$

or, $$r = \frac{mv^2}{F} = \frac{mv^2}{ma} = \frac{v^2}{a} = \frac{(3\times10^5\ m/s)^2}{1.81\times10^{12}\ m/s^2} = 49.7\ mm.$$

***Example 7.4** A proton ($q = +e$, $m = 1.67\times10^{-27}$ kg) is shot with speed 8×10^6 m/s at an angle of 30° to an x-directed field B = 0.15 T. Describe the path followed by the proton.*

Soln.

Let the velocity of the proton be resolved into components

parallel to and perpendicular to the magnetic field.

$$v_{||} = v\cos\theta = (0.86 \times 8 \times 10^6 \text{ m/s}) = 6.88 \times 10^6 \text{m/s}$$

$$v_{\perp} = v\sin\theta = (0.5)(8 \times 10^6 \text{m/s}) = 4 \times 10^6 \text{m/s}.$$

The magnetic force due to $v_{||}$ is zero; since $\sin\theta = 0$; the magnetic force due to $v_{\perp}$ has no x-component. Therefore the x-motion is uniform, at speed $v_{||} = 6.88 \times 10^6$ m/s while the transverse motion is circular with radius.

$$r = \frac{mv_{\perp}}{qB} = \frac{(1.67\times10^{-27}\,\text{kg})(0.5\times8\times10^6\,\text{m/s})}{(1.6\times10^{-19}\text{C})(0.15\text{T})}$$

$$= 0.28 \text{ m}.$$

The proton will spiral along the x-axis; the radius of the spiral (or helix) will be 28 cm.

The pitch of the helix is defined as the distance traveled by the proton along the x-axis during one revolution. Now the time taken to complete one revolution is

$$T = \frac{2\pi r}{v_{\perp}} = \frac{2\pi(0.28\text{m})}{4\times10^6\,\text{m/s}} = 4.4 \times 10^{-7}\text{s}.$$

During that time, the proton will travel an x-distance of

$$\text{pitch} = (v_{||})(\text{period}) = (6.88 \times 10^6 \text{ m/s})(4.4 \times 10^{-7}\text{s})$$

$$= 3.0\text{m}.$$

***Example 7.5** (a) Show that the pitch p of the helical path of a charged particle moving with velocity v at an angle θ with the magnetic field B is 2πmvcosθ/qB, where q in the charge and m is the mass of the particle. (b) Suppose the particle in (a) is a 2KeV electron projected into the field B of 0.20 W/m² with its velocity making an angle θ=86°. Find the period T, the pitch p and the radius r of the helix.*

Sole.

(a) The velocity of the particle along the axis of the helix *i.e.* in a direction parallel to **B** is $v\cos\theta$ and the velocity perpendicular to **B** is $v\sin\theta$. It is the perpendicular component which is responsible for circular motion. The period of circular motion is given by

$$T = \frac{2\pi r}{v\sin\theta} = \frac{2\pi m}{qB}$$

Hence the pitch of the helix *i.e.*, the distance traversed along the axis during the time the particle makes one rotation, is

$$p = v\cos\theta.T = 2\pi mv\cos\theta/qB$$

(b) $$T = \frac{2\pi m}{qB} = \frac{(2)(3.141)(9.1\times10^{-31}kg)}{(1.6\times10^{-19}C)(0.20Wb/m^2)}$$

$$= 1.8 \times 10^{-10} \text{ sec.}$$

Kinetic energy $K = (2 \times 10^3 eV)(1.6 \times 10^{-19} J/eV)$

$$= 3.2 \times 10^{-16} \text{ Joules.}$$

$$v = \sqrt{\frac{2K}{m}} = \sqrt{\frac{(2)(3.2\times10^{-16}J)}{(9.1\times10^{-31}kg)}} = 2.65 \times 10^7 \text{ m/sec.}$$

Hence,

$$p = Tv\cos\theta = (1.8 \times 10^{-10}s)(2.65 \times 10^7 m/s)(\cos 86°)$$

$$= (1.8 \times 10^{-10}s)(2.65 \times 10^7 m/s)(0.070)$$

$$= 0.33 \text{ mm.}$$

Now,

$$r = \frac{mv\sin\theta}{qB} = \frac{(9.1\times10^{-31}kg)(2.65\times10^7 m/s)(\sin 86°)}{(1.6\times10^{-19}C)(0.20Wb/m^2)}$$

$$= \frac{(9.1\times10^{-31}kg)(2.65\times10^7 m/s)(0.998)}{(1.6\times10^{-19}C)(0.20Wb/m^2)}$$

$$= 0.75 \text{ mm.}$$

Example 7.6 *After being accelerated though a potential difference of 5000 V, a singly charged carbon ion moves in a circle of radius 21 cm in the magnetic field of a mass spectrometer. What is the magnitude of the field ?*

Soln.

After being accelerated through a potential difference of 5000 V, the kinetic energy of singly charged carbon ion is

$$\frac{1}{2} m_C v^2 = qV \; ; \qquad \text{or,} \quad v = \sqrt{\frac{2qV}{m_C}}$$

where m_C = mass of the carbon ion

$$= (12u)(1.66 \times 10^{-27} \text{ kg/u}) = 19.9 \times 10^{-27} \text{kg}.$$

Again, we have

$$\frac{m_C v^2}{r} = qvB$$

$$\text{or,} \quad B = \frac{m_C v^2}{qvr} = \frac{m_C v}{qr} = \frac{m_C}{qr}\sqrt{\frac{2qV}{m_C}}$$

$$\text{or,} \quad B^2 = \frac{m_C^2}{q^2 r^2}\frac{2qV}{m_C} = \frac{2Vm_C}{r^2 q}$$

$$= \frac{(2)(5000\text{V})(19.9 \times 10^{-27} \text{kg})}{(0.21\text{m})^2 (1.6 \times 10^{-19} \text{C})}$$

$$= 2.82 \times 10^{-2} \text{ T}^2$$

$$\therefore \; B = 0.168 \text{ T}.$$

The parallel component determines the pitch of the helix which is defined as the x-distance travelled during one revolution. Or

$$\text{pitch} = v_{11} \times T$$

$$= v_{11} \times \frac{2\pi r}{v_1}$$

$$= v \cos\theta \times \frac{2\pi r}{v \sin\theta}$$

Example 7.7 A beam of electrons with kinetic energy 22.5 eV circulate around a circular path-in a uniform magnetic field of 4.55 x 10^4 T. (a) what is the radius of the electron's path? (b) what is the frequency of the circulating electrons? (c) what is the period of revolution.

Soln.

(a) Radius r = mv/qB;

The velocity of the electrons is given by

$$V = \sqrt{\frac{2K}{m}} = \sqrt{\frac{(2)(22.5\times1.6\times10^{-19}\,J)}{9.1\times10^{-31}\,kg}} = 2.81 \times 10^6 \text{ m/s}$$

$$\therefore\ r = \frac{mv}{qB} = \frac{(9.11\times10^{-31}\,kg)(2.81\times10^{6}\,m/s)}{(1.6\times10^{-19}\,C)(4.55\times10^{-4}\,T)} = 3.52 \text{ cm}$$

(b) The frequency of the circulating electrons

$$f = \frac{qB}{2\pi m} = \frac{(1.60\times10^{-19}\,C)(4.55\times10^{-4}\,T)}{2\pi(9.11\times10^{-31}\,kg)}$$

$$= 1.27 \times 10^7 \text{ Hz} = 12.7 \text{ MHz.}$$

(c) time period of revolution

$$T = \frac{1}{f} = \frac{1}{12.7\times10^{6}\,Hz} = 7.86 \times 10^{-8}\text{s} = 78.6 \text{ ns.}$$

7.4 Magnetic force on a current carrying conductor (Tapon)

A current is an assembly of moving charges. Since a magnetic field exerts a sideways force on a moving charge, we expect that it will also exert a sideways force on a wire carrying a current. (Let us consider a length l of a wire carrying a current i and placed in a field of magnetic induction **B** as shown in Fig 7.2. For simplicity the wire is so oriented that the current density vector **j** is at right angles to **B**.

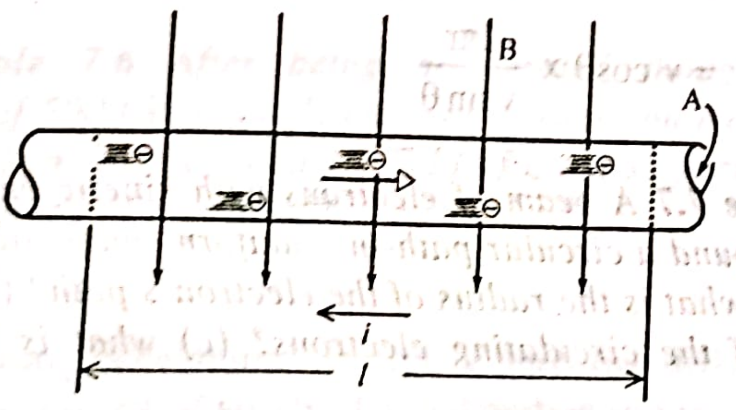


Fig. 7.2

The current i in a metal wire is carried by the free (or conduction) electrons. Let n be the number of free electrons per unit volume of the wire. The magnitude of the average force experienced by one such electron is given by

$$F' = q_0 vB \sin\theta$$

$$= ev_d B.$$

where e (= q_0), the charge on one electron, v_d (= v) is the drift speed of the electrons and $\theta = 90^\circ$. From the relation $v_d = j/ne$ (eqn. 5.6), we have

$$F' = e\left(\frac{j}{ne}\right)B = \frac{jB}{n}$$

If A is the area of cross-section of the wire, then the volume of the wire of length l is Al. Hence the total number of free electrons in the wire is nAl. Then the total force acting on the free elections in the wire and hence on the wire itself is

$$F = F'(nAl) = nAl \cdot \frac{jB}{n} = AljB \qquad (7.11)$$

But jA (from $j = \frac{i}{A}$) is the current i in the wire. Hence eqn. 7.11

reduces to

$$F = ilB \qquad (7.12)$$

The negative charges, which move to the right in the wire of Fig. 7.2, are equivalent to positive charges moving to the left, *i.e.*, in the direction of the central arrow. The velocity **v** would then point to the left and the force on the wire as given by eqn. 7.2 points up, out of the page. It the charge carriers actually are negative charges, then **v** will point to the right, but since q_0 is negative, the direction of the force will be same as before. Thus by measuring the sideways magnetic force on a wire carrying a current and placed in a magnetic field, it cannot be ascertained if the charge carriers are positive or negative charges.

Eqn. 7.12 holds only if the wire is at right angles to **B**. The more general vector form of the equation will be

$$\mathbf{F} = i\,\mathbf{l} \times \mathbf{B} \qquad (7.13)$$

where l is a length vector that points along the (straight) wire in the direction of the current.

Eqn. 7.13 and eqn. 7.2 are equivalent to one another in the sense that either can be taken as a defining equation for **B**. But since, it is much easier to measure the magnetic force acting on a wire than on a single moving change, **B**, in practice, is defined from eqn. 7.13.

It the wire is not straight, we can imagine it broken up into small straight segments. Considering one such segment of length **dl**, the force **dF** acting on it can be given by

$$\mathbf{dF} = i\mathbf{dl} \times \mathbf{B} \tag{7.14}$$

By integrating eqn. 7.14 in an appropriate manner, the force **F** acting on a non-linear conduction can be found. As an example let us consider that a wire carrying a current i and bent as shown in Fig. 7.3 is placed in a uniform field of magnetic induction **B** that emerges from the plane of the figure. We would like to calculate the force acting on the wire.

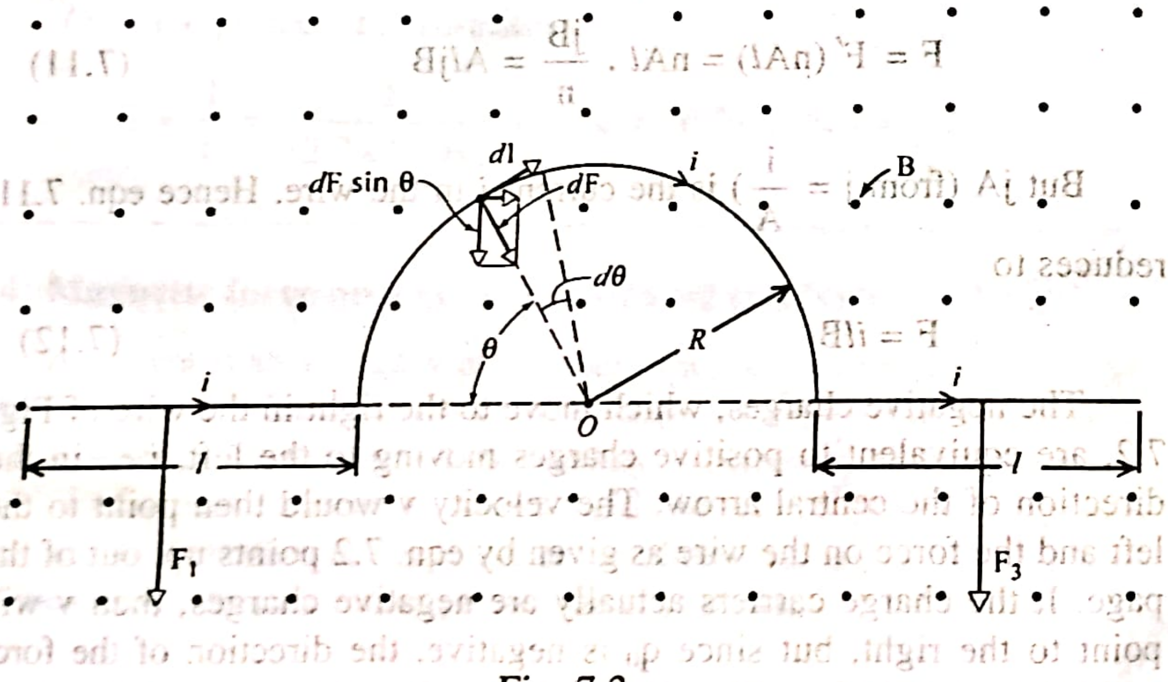


Fig. 7.3

The magnetic field is represented by lines of induction, shown emerging from the page. The dots show that the sense of **B** *is up, out of the page.*

The magnitude of the force that acts on each straight segment of length l is given by eqn. 7.13.

$$F_1 = F_3 = i\,l\,B$$

and points as shown by the arrows in the figure.

Let us consider a segment of the central arc of length dl. The force dF acting on this segment is given by

$$dF = iB\,dl = iB\,(Rd\theta)$$

where R is the radius of the arc and $d\theta$ is the angle subtended by the segment dl at the centre O of the arc. The direction of dF is radially toward O. It may be noticed that only the downward component $dF \sin\theta$ of the force is effective, the horizontal component $dF\cos\theta$ being cancelled by an oppositely directed component associated with a symmetrically located segment on the other side of O. Thus the total force on the central arc (a semicircle of wire of radius R) points down and is

$$F_2 = \int_0^{\pi} dF \sin\theta = \int_0^{\pi} (iBRd\theta)\sin\theta$$

$$= iBR\int_0^{\pi} \sin\theta\, d\theta = 2iBR$$

The resultant force on the whole wire is then

$$F = F_1 + F_2 + F_3 = ilB + 2iBR + ilB$$

$$= 2iB\,(l + R)$$

It may be noticed that this force is the same as that acting on a straight wire of length $(2l+2R)$

Example 7.8 *At the equator, the earth's magnetic field is nearly horizontal, directed from the southern to northern hemisphere. Its magnitude is about 0.50G. Find the force (direction and magnitude) on a 20-m wire carrying a current of 30A parallel to the earth (a) from east to west, (b) from north to south.*

Soln.

(a) $F = ilB \sin\theta$

$= (30)\,(20)\,(5 \times 10^{-5}) \sin 90°$

$= (30)\,(20)\,(5 \times 10^{-5})\,(1)$

$= 0.030$ N.

$10^4 G = 1T$

$0.5G = 5 \times 10^{-5}$ T

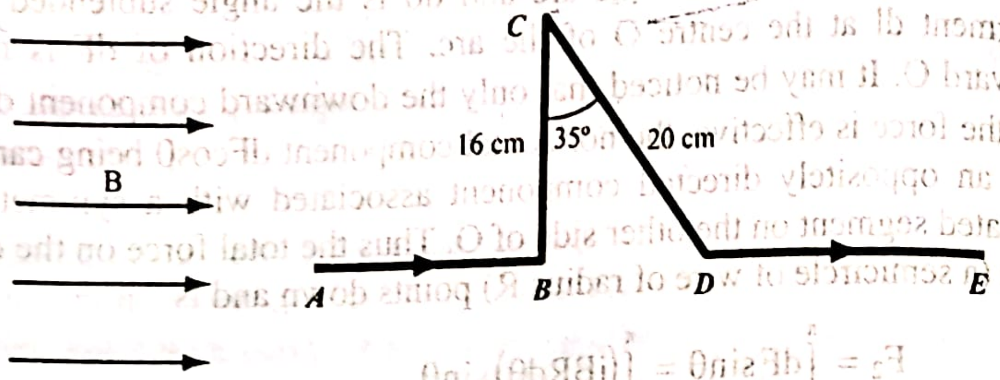
$\theta = 90°$

(b) $F = ilB \sin\theta$ $\theta = 0°$

$= (30)(20)(5 \times 10^{-5}) \sin 0°$

$= 0.$

Example 7.9 *Find the force on each segment of the wire shown in the figure below, if B= 0.15T. Assume that the current in the wire is 5A.*



Soln.

For the straight segments AB and DE, l and B are parallel ($\theta = 0°$); so $F = 0$.

In section BC, $\theta = 90°$, $l = 16cm = 0.16m$.

$F = ilB \sin 90° = (5)(0.16)(0.15)(1) = 0.12N$, into page.

For the section CD, $\theta = 65°$, $l = 20cm = 0.20m$.

$F = (5)(0.20)(0.15) \sin 65° = (5)(0.20)(0.15) = 0.136N$, out of page.

Example 7.10 A straight horizontal segment of copper wire carries a current i = 28A. What are the magnitude and direction of the magnetic field needed to float the wire, that is, to balance its weight? Its linear mass density is 46.6g/m

Soln.

For a length l of the wire we have

$$mg = F = ilB$$

$$\text{or, } B = \frac{mg}{il} = \frac{(m/l)g}{i} = \frac{(46.6\times10^{-3}\text{kg/m})(9.8\text{m/s}^2)}{28\text{A}}$$

$$= 1.6 \times 10^{-2}\ T = 16mT.$$

7.5 Torque on a current loop

When a loop of wire carrying a current i is placed in a magnetic field, that loop can experience a torque tending to rotate it about a particular axis, generally taken to pass through the centre of mass of the loop. This principle is the basis of operation of electric motors as well as galvanometers on which analog current and voltage metres are based.

Fig. 7.4 shows a rectangular loop of sides a and b carrying a current i in a uniform magnetic field **B**. For simplicity, only the loop itself is shown. It is assumed that the wires that bring current to and from the loop are twisted together so that there is no net magnetic force on them. It is further assumed that the loop is suspended in such a way that it is free to rotate about any axis.

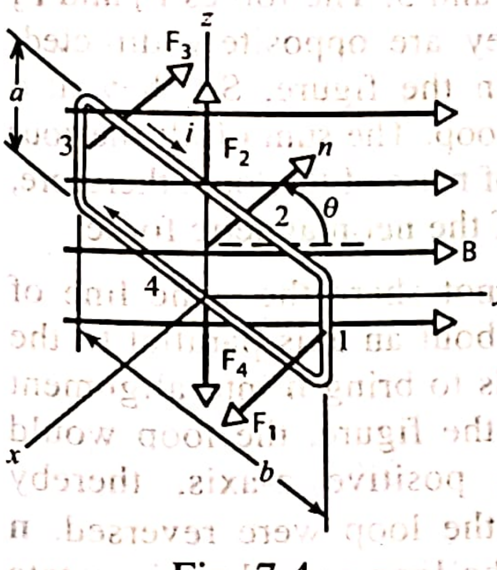


Fig. 7.4

The uniform field **B** is in the y-direction of the coordinate system of Fig. 7.4. The loop is so oriented that the Z–axis lies in its plane and its sides 1 and 3 are perpendicular to **B**. The plane of the loop is indicated by a unit vector **n** that is perpendicular to the plane. [*A unit vector is a vector that has a magnitude of exactly 1 and that points in a particular direction. It lacks both dimension and unit. Its sole purpose is to point, that is, to specify a direction*]

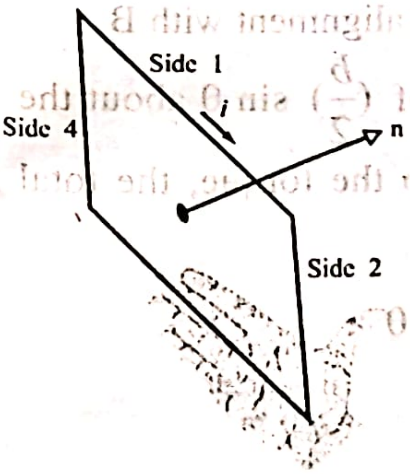


Fig. 7.5

The direction of the unit vector **n** is determined by using the right hand rule, so that if the fingers of your right hand are curled or pointed in the direction of the current at any point in the loop, the extended thumb gives the direction of **n** (Fig. 7.5). Let the angle made by the unit vector n with B be θ.

The net torque on the loop can be determined by calculating the force on each of its four sides, using the relation $\mathbf{F} = i\mathbf{l} \times \mathbf{B}$

remembering that the force on each segment must be perpendicular to both B and to the direction of current in the segment. Thus the magnitude of the force F_2 on side 2 (of length b) is

$$F_2 \; ibB \sin (90° - \theta) = ibB \cos \theta$$

where $(90° - \theta)$ is the angle between the vectors l and **B**. This force points in the positive Z - direction. The force F_4 on side 4 has the magnitude

$$F_4 = ibB \sin (90° + \theta) = ibB \cos \theta$$

and points in the negative Z–direction.

$\mathbf{F_2}$ and $\mathbf{F_4}$ are equal and opposite; they cancel out exactly and so contribute nothing to the net force on the loop. Furthermore, they have the same line of action, so the net torque due to these two forces is also zero.

The situation is different for sides 1 and 3. The forces F_1 and F_3 have a common magnitude of iaB. They are oppositely directed, parallel and antiparallel to the x–axis in the figure. So they also contribute nothing to the net force on the loop. The sum of all the four forces gives a resultant of zero; the centre of mass of the loop, therefore, does not accelerate under the influence of the net magnetic force.

However, the forces $\mathbf{F_1}$ and $\mathbf{F_3}$ do not share the same line of action; so they do tend to turn the coil about an axis parallel to the z–axis. The direction of the rotation tends to bring **n** into alignment with B. Thus in the situation shown in the figure, the loop would rotate clockwise as viewed from the positive z-axis, thereby reducing the angle θ. If the current in the loop were reversed, **n** would have the opposite direction, and the loop would again rotate through the angle $(\pi - \theta)$ necessary to bring **n** into alignment with B.

The forces $\mathbf{F_1}$ and $\mathbf{F_3}$ have moment arms of $(\frac{b}{2}) \sin \theta$ about the z–axis. Since both forces contribute equally to the torque, the total torque on the loop is

$$\tau' = (ia\,B)\left(\frac{b}{2}\right) \sin \theta + (ia\,B)\left(\frac{b}{2}\right) \sin \theta$$

$$= 2\,(ia\,B)\,\frac{b}{2}\,\sin \theta$$

$$= iab\,B \sin \theta \qquad (7.15)$$

Note that if **n** is already parallel to **B**, *i.e.*, $\theta = 0$, then there is no torque.

Eqn. 7.15 gives the torque on a single rectangular loop in the field. If the loop (or the coil) consists of N turns, such as might be found in a motor or a galvanometer, then eqn. 7.15 gives the torque on each turn. The total torque on the loop will then be given by

$$\tau = N.\tau' = NiabB \sin\theta$$

$$= (NiA)\, B \sin\theta \qquad (7.16)$$

where A = ab is the area of the rectangular loop.

The quantities (NiA) are grouped together because they are all properties of the coil: its number of turns, its area and the current it carries. The equation holds for *all plane loops of area A, whether they are rectangular or not.*

7.6 The Magnetic Dipole

In chapter I, we have discussed the effect of an electric field **E** on an electric dipole, which consists of two equal and opposite charges separated by a small distance. We have found that the effect of the electric field was to exert a torque on the electric dipole that tended to rotate the dipole so that the dipole moment **p** is aligned with **E**. As outlined in the previous section, the effect of a magnetic field on a current loop is to produce a torque on the loop which tends to rotate it so that the normal vector **n** aligns with **B**. This similarity suggests that equations similar to those for the electric dipole may be used to analyze the effect of magnetic field on a current loop. This analogy is further encouraged by the similarity of the magnetic field lines of a bar magnet, which is an example of a *magnetic dipole.*

The torque of an electric dipole is given by

$$\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}$$

which can also be written in terms of magnitude as

$$\tau = pE \sin\theta$$

where θ is the angle between **p** and **E**. The torque produced on a

current carrying loop, as derived in the previous section is

$$\tau = NiA\ B \sin\theta$$

The similarity between the two expressions is striking. In each case the appropriate field (E or B) appears, as does a term sinθ. Comparison suggests the NiA in eqn. 7.16 can be taken as the *magnetic dipole moment* μ, corresponding to **p** in eqn. $\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}$, or

$$\mu = NiA$$

The direction of the magnetic dipole moment must be parallel to **n** *i.e.*, with the fingers of the right hand in the direction of the current, the extended thumb gives the direction of μ. Eqn. 7.16 can, therefore, be written in the vector form, as

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} \qquad (7.17)$$

It can be shown that eqn. 7.17 in general, gives the most general description of the torque exerted on *any* planar current loop in a uniform magnetic field **B**. It holds no matter what the shape of the loop or the angle between its plane and the field.

The analogy can be extended further. Since a torque acts on a current loop or other magnetic dipole, when it is placed in an external magnetic field, it follows that work (positive or negative) must be done by an external agent to change the orientation of such a dipole. The magnetic dipole must then have a *magnetic potential energy* that depends on the orientation of the dipole in the field. For an electric dipole, the expression for the potential energy is given by

$$U(\theta) = -\mathbf{p}.\mathbf{E}$$

In strict analogy, the expression for the potential energy of magnetic dipole can be written as

$$U(\theta) = -\boldsymbol{\mu}.\mathbf{B}$$

By analogy with the assumption made for electric dipoles in chapter I, it is assumed that the magnetic energy U is zero when μ and B are at right angles, that is when θ = 90°. This choice of a zero - energy configuration for U is arbitrary as we are interested only in the *changes* in energy that occur when the dipole is rotated.

Thus a magnetic dipole has its lowest energy (= − μB cos0° =

$-\mu B$) when its dipole moment is lined up with the magnetic field. It has its greatest energy ($= -\mu B \cos 180° = +\mu B$) when it points in a direction opposite to the field. The difference in energy between the two orientations is

$$\Delta U = U(180°) - U(0°)$$

$$= (-\mu B \cos 180°) - (-\mu B \cos 0°)$$

$$= +\mu B - (-\mu B)$$

$$= 2\mu B. \quad (7.18)$$

Eqn. 7.18 gives the work that must be done by an external agent to turn a magnetic dipole through 180^0 starting form its initial position when it was lined up with the magnetic field.

***Example 7.11** A coil of 20 turns has an area of 800 mm² and bears a current of 0.5A. It is placed with its plane parallel to magnetic field of intensity 0.3T. Determine the torque on the coil.*

Soln.

$$\tau = NiAB \sin\theta$$

$$= (20)(0.5)(800 \times 10^{-6})(0.3)(1)$$

$$= 2.4 \times 10^{-3} \text{ N.m.}$$

$N = 20$
$i = 0.5A$
$A = 800 \text{ mm}^2 = 800 \times 10^{-6} \text{m}^2$
$\theta = 90°$, $B = 0.3T$

***Example 7.12** A rectangular loop 6 cm high and 2 cm wide is placed in a magnetic field of 0.02T. If the loop contains 200 turns and carries a current of 50 mA, what is the torque on it? Assume that the face of the loop is parallel to the field.*

Soln.

$$\tau = NiAB \sin\theta$$

$$= (200)(50 \times 10^{-3})(12 \times 10^{-4})(0.02)(1)$$

$$= 2.4 \times 10^{-4} \text{ N.m.}$$

$A = 6 \times 2 \times 10^{-4} \text{m}^2$
$N = 200$
$i = 50 \times 10^{-3} A$
$B = 0.02T$, $\theta = 90°$

Example 7.13 *Analog voltmeters and ammeters, in which the reading is displayed by the deflection of a pointer over a scale, work by measuring the torque exerted by a magnetic field on a current loop. A spring provides a counter torque that balances the magnetic torque, resulting in a steady angular deflection ϕ corresponding to a given steady current i in the coil. The coil is 2.1 cm high and 1.2 cm wide; it has 250 turns and is mounted so that it can rotate about its axis in a uniform radial magnetic field with B = 0.23T. If a current of 100 μA produces an angular deflection of 28° (= 0.49 rad) what must be the torsional constant k of the spring?*

Soln.

Since the magnetic torque is equal to the restoring torque $k\varphi$ of the spring

$$\tau = NiAB \sin\theta = k\phi$$

Here $N = 250$, $i = 100 \times 10^{-6}$A, $B = 0.23$T, $\theta = 90°$,

$A = (2.1 \times 1.2 \times 10^{-4}\text{m}^2)$ and $\phi = 0.49$ rad.

$$\therefore\ (250)(100 \times 10^{-6})(2.52 \times 10^{-4})(0.23)(1) = k(0.49)$$

$$\text{or, } k = \frac{(250)(100\times10^{-6})(2.52\times10^{-4})(0.23)}{0.49}$$

$$= 3.0 \times 10^{-6}\text{N.m/rad.}$$

Example 7.14 *A galvanometer, which is to have an internal resistance of 1.00 Ω, is wound with a copper wire whose resistance is 33.86 Ω/km. The dimensions of the rectangular coil are 2.50 cm by 2.00 cm (a) How many turns will the coil have? (b) The magnetic field in the galvanometer is 0.40T. The spring constant k equals 5.00 $\times 10^{-6}$N.m/rad. The galvanometer gives a full-scale deflection with a current of 1.00 mA. What is the full-scale deflection angle in radians? The plane of the coil is parallel to the field.*

Soln.

(a) The length of a 1.00 Ω piece of copper wire is 1 Ω / (338.6Ω/km) = ($10^{5.0}$/338.6) cm = 295.3cm. The perimeter of the coil is 2(2.50 + 2.00) = 9.00 cm. So the required number of turns N = 295.3/9.00=32.8 ≃ 33turns.

(b) We know that

$$\tau = NiABsin\theta = k\phi$$

where $N = 33$, $i = 1 \times 10^{-3}A$, $A = (2.50 \times 2.00 \times 10^{-4}m^2)$

$B = 0.40\ T$, $\theta = 90°$, $k = 5.00 \times 10^{-6}N.m/rad.$

$$\therefore \quad \phi = \frac{NiABsin\theta}{k} = \frac{(33)(1\times10^{-3})(5\times10^{-4})(0.40)(1)}{5.00\times10^{-6}}$$

$$= 1.32\ rad = 75.6° \simeq 76°.$$

***Example 7.15** A coil having an area of 2.52 cm^2 and 250 turns carries a current of 85 μA. (a) What is the magnetic dipole moment of the coil? (b) The magnetic dipole moment of the coil is lined with an external magnetic field whose strength is 0.85T. How much work would be done by an external agent to rotate the coil through 180°?*

Soln.

(a) The magnitude of the magnetic dipole moment,

$$\mu = NiA = (250)(85 \times 10^{-6}A)(2.52 \times 10^{-4}m^2)$$

$$= 5.36 \times 10^{-6}\ A.m^2$$

$$= 5.36 \times 10^{-6}\ J/T.$$

(b) The external work is equal to the increase in potential energy of the system, which is .

$$W = \Delta U = -\mu B \cos 180° - (-\mu B \cos 0°)$$

$$= 2\mu B$$

$$= (2)(5.36 \times 10^{-6}\ J/T)(0.85T)$$

$$= 9.1 \times 10^{-6}\ J = 9.1\ \mu J.$$

7.7 The Hall Effect

(Tapon)

In 1879, Edwin H. Hall, then a 24 year old research student at the Johns Hopkins University, conducted an experiment that permitted direct measurement of the sign and the number density (number per unit volume) of charge carriers in a conductor.

Consider a flat strip of material of width w carrying a current i, as shown in Fig. 7.6. As usual, the direction of the current arrow, labeled i, is the direction in which the charge carriers would move *if* they were positive, or opposite to the direction in which the negative charge carriers would move. A uniform magnetic field **B** is established

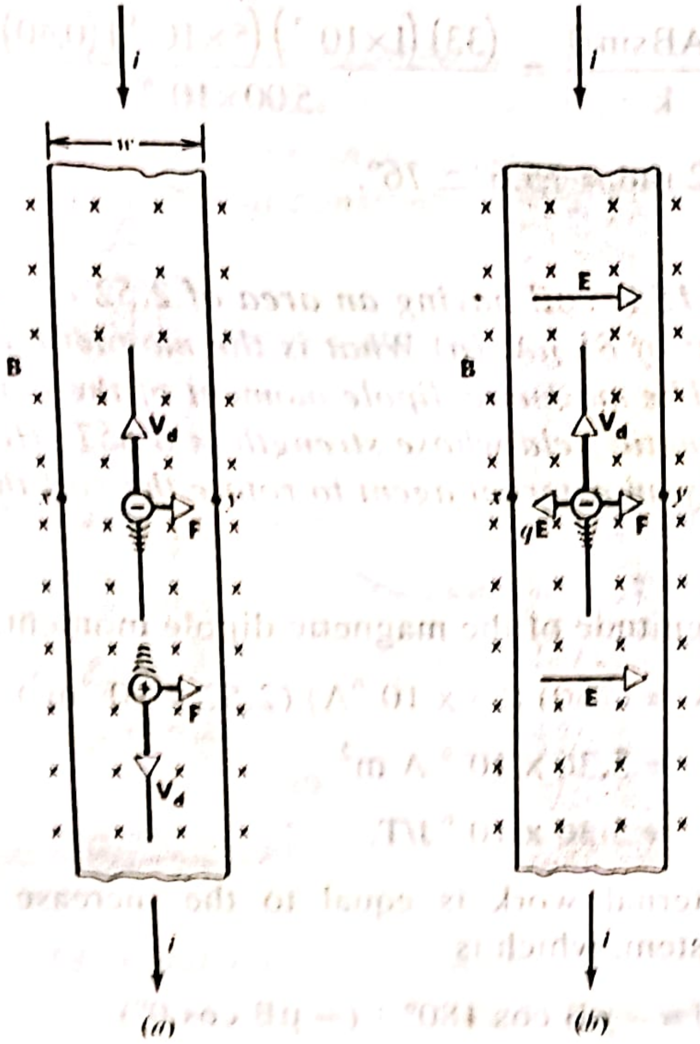


Fig. 7.6

perpendicular to the plane of the strip, such as by placing the strip between the poles of an electromagnet. This field exerts a deflecting force F on the strip (given by i **l** × **B**) which points to the right of the figure. Since the sideways force on the strip is due to the sideways forces on the charge carriers (given by q **v** × **B**), it follows that these carriers, whether they are positive or negative, will tend to drift toward the right in Fig. 7.6 as they drift along the strip. The buildup of charge along the right side of the strip (and a corresponding deficiency of charge of that sign on the opposite side of the strip)

produces a transverse potential difference V_{xy} between points such as x and y. This potential difference is called the *Hall potential difference.* V_{xy} can be measured by connecting the leads of a voltmeter to points x and y. If the carriers are positive y will be at a higher potential than x. Thus the sign of V_{xy} gives the sign of the charge carriers and the magnitude of V_{xy} gives their density (number per unit volume) Experiment shows that in metals the charge carriers are negative. This may seem like an obvious conclusion today; however, it should be kept in mind that Hall's work was done nearly 20 years before Thomson's discovery of the electron, and the nature of electrical conduction in metals was not at all obvious at that time.

The charge carriers do not build up without limit on the right edge of the strip. Let us assume that conduction in the material is due to charge carriers of a particular sign (positive or negative) moving with drift velocity $\mathbf{v_d}$. As the charge carriers drift, they are deflected to the right in Fig. 7.6 by the magnetic force. As the charges collect on the right side, they set up a transverse Hall electric field that acts inside the conductor to oppose the sideway drift of additional charge carriers. This Hall electric field is another manifestation of the Hall potential difference and is related to it by

$$E_H = V_{xy} / w$$

Eventually an equilibrium is reached in which the sidesways magnetic deflecting force on the charge carriers ($q\,\mathbf{v_d} \times \mathbf{B}$) is just cancelled by the oppositely directed electric force ($q\,\mathbf{E_H}$) caused by the Hall electric field. In vector terms, the Lorentz force on the charge carriers under these circumstances is zero, or

$$q\mathbf{E}_H + q\mathbf{v}_d \times \mathbf{B} = 0$$

$$\text{or,} \qquad \mathbf{E}_H = -\mathbf{v}_d \times \mathbf{B} \tag{7.19}$$

This equation shows explicitly that if $\mathbf{E_H}$ and $\mathbf{B}$ are measured $\mathbf{v_d}$ can be determined both in magnitude and direction; given the direction of $\mathbf{v_d}$, the sign of the charge carriers follows at once, as Fig. 7.6 shows.

The number of charge carriers per unit volume (n) can also be found from Hall effect measurements. Since $\mathbf{v_d}$ and $\mathbf{B}$ are at right

angles, we can write eqn. 7.18 in terms of magnitude as

$$E_H = v_d B.$$

Form eqn. 5.6 chapter V, we can write the drift speed as $v_d = j/ne$, where j in the current density in the strip and n is the number of charge carriers per unit volume. The current density j is the current i per unit cross-sectional area A of the strip. If t is the thickness of the strip, then its cross-sectional area A can be written as wt. Substituting V_{xy}/w for the electric field E_H, we obtain

$$\frac{V_{xy}}{w} = v_d B = \frac{j}{ne} B = \frac{i}{wtne} B$$

or, solving for the density of charge carriers,

$$n = \frac{iB}{etV_{xy}} \qquad (7.20)$$

From a measurement of the magnitude of the Hall potential difference V_{xy} we can find the density of the charge carriers.

***Example 7.16** A strip of copper 150μm thick is placed in magnetic field B of magnitude 0.65 T, and a current i = 23A is sent through the strip. What Hall potential difference V will appear across the width of the strip? The number of charge carriers per unit volume of copper is $8.49 \times 10^{28}\ m^{-3}$.*

Soln.

From eqn. 7.19 ($n = \frac{iB}{etV}$)

we get

$$V = \frac{iB}{net} = \frac{(23)(0.65)}{(8.49\times10^{28})(1.6\times10^{-19})(150\times10^{-6})}$$

$$= 7.4 \times 10^{-6} V$$

$$= 7.4\ \mu V.$$

EXERCISES

1. A charged particle is moving through a magnetic field of magnetic induction B. Write down the equation defining magnetic induction in terms of the measurable quantities and explain the characteristics of the force acting on the particle.

2. Of the three vectors in the equation $\mathbf{F} = q\,\mathbf{v} \times \mathbf{B}$, which pairs are always at right angles? Which may have any angle between them?

3. Why is the magnetic force acting on a charged particle is always a sideways force? How would you determine the direction of this force? What is the unit of force?

4. A current carrying conductor of a given length is placed in a magnetic field. Obtain an expression for the force exerted on the wire.

5. A charged particle is projected into a magnetic field in a direction perpendicular to the field. Show that the particle moves in a circular path. Obtain expressions for the radius of the circular path and the period for one revolution. Does the period depend on the velocity of the particle?

6. A charged particle is projected into a magnetic field in a direction making an angle with the direction of magnetic field. Show that the particle will move about the direction of the field in a helical path. What is the pitch of the helix?

7. Show that a current carrying loop immersed in a magnetic field is subjected to a torque which tends to rotate the loop around its axis. Hence obtain an expression for the torque acting on the loop.

8. Obtain an expression for the torque acting on a current carrying coil placed in a magnetic field.

9. Explain the terms magnetic dipole and magnetic dipole moment.

10. A magnetic dipole is initially lined up with the magnetic field. Obtain an expression for the work that must be done by an external agent to turn the magnetic dipole through 180^0.

11. What is Hall effect? Show how Hall effect can be used to determine the nature of the charge carriers and the number of charge carriers per unit volume of a conductor.

12. An ion ($q = +2e$) enters a magnetic field with flux density 1.2 Wb/m^2 at velocity of 2.5×10^5 m/s perpendicular to the field. Determine the force on the ion. **[9.6×10^{-14} N]**

13. An electron is accelerated form rest through a potential difference of 3750V. It enters a region where B = 4 mT perpendicular to its velocity. Calculate the radius of the path it will follow. **[52 mm]**

14. What might be the mass of a positive ion that is moving at 10^7 m/s and is bent into a circular path of radius 1.55m by a magnetic field of 0.134 Wb/m^2. The charge on the ion may be assumed to be q = ne where n is a positive integer. **[(3.3×10^{-27} kg) n]**

15. A 10-ev electron is circulating in a plane at right angles to a uniform field of magnetic induction of 1.0×10^{-4} Weber/ metre^2. (a) What is the radius of its orbit? (b) What is the period of revolution? (c) What is the frequency of revolution? **[11cm, 3.6×10^{-7} see, 2.8 × 10^6 rev/sec]**

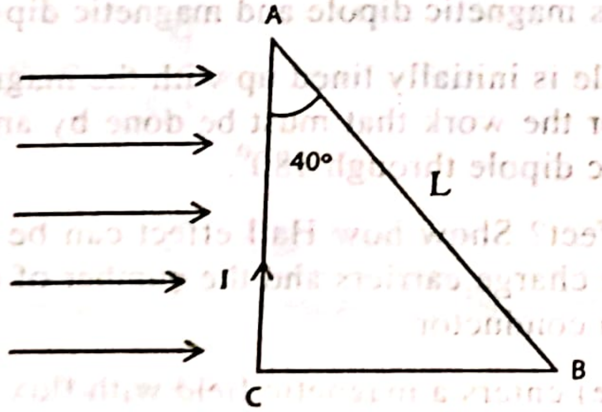
16. 2-KeV alpha particles (m = 6.68 × 10^{-27} kg, q = +2e) enter a magnetic field B = 0.2T perpendicular to their direction of motion. Calculate the radius of their path. **[32mm]**

17. A cathode-ray beam (an electron beam; m = 9.1 × 10^{-31} kg q = –e) is bent in a circle of radius 2 cm by a uniform field with B = 4.5mT. What is the speed of the electrons? **[1.58×10^{7}m/s]**

18. An electron is shot with speed 5×10^{6}m/s out from the origin of coordinates. Its initial velocity makes an angle of 20^0 to the +x axis. Describe its motion if a magnetic field B = 2.0 mT exists in the +x direction. **[electron follows a helical path of radius 0.48cm and pitch 8.33 cm]**

19. A wire bearing a current of 10A lies perpendicular to a uniform magnetic field. A force of 0.2N is found to exist on a section of the wire 80 cm long. Determine the magnetic induction B. **[0.025T]**

20. Find the magnitude and direction of the force on wire (a) AC, (b) BC, (c) AB and (d) ABCA:



[(a) 0.77 iLB into the page (b) 0 (c) 0.77 iLB out of the page (d) 0.77 iLB in + 0 + 0.77 iLB out = 0]

21. A rectangular loop of sides 12 cm and 10 cm and of 40 turns carries a current of 2A in a magnetic field B = 0.25T, the plane of the loop being parallel to the direction of magnetic filed. Find the torque on it **[0.24 N.m]**

22. A coil is 2.1 cm high and 1.2 cm wide and has 250 turns. The coil is mounted so that it can rotate about an axis in a uniform radial magnetic field with B = 0.85T. If the coil carries a current of 100 μA, (a) what is the magnitude of the magnetic dipole moment of the coil? (b) how much work would be required to turn the coil end for end? **[6.3×10^{-6} A.m^2, $\simeq$ 11μ J]**

23. A copper strip 2.0 cm wide and 1.0 mm thick is placed in magnetic field with B = 1.5 Webers/metre^2. If a current of 200 amp is set up in the strip, what Hall potential difference appears across the strip? The number of charge carriers per unit volume of copper is 8.4×10^{28}/m^3. **[22μv]**

CHAPTER - VIII

MAGNETIC EFFECTS OF CURRENTS

8.1 The Biot–Savart Law

The forces exerted by a magnetic field **B** on a moving charge or a current carrying conductor and the torque produced on a magnetic dipole have been discussed in the previous chapter. The magnetic effects produced by a current or by moving charges will now be discussed.

The fact that currents produce magnetic fields was first discovered by Oersted in 1820. He was able to demonstrate this by bringing in a compass needle near a current carrying conductor when the needle was deflected. It was further demonstrated that the lines of induction of the field produced by the current form closed circles around the wire. Just after a few weeks of this discovery, Ampere presented the results of a series of experiments showing that two current carrying conductors exert force on each other. This was explained by considering, in the light of Oersted's experiment, that one conductor produces a magnetic field at the site of the second conductor and exerts a force on it.

We shall now develop a procedure for calculating the magnetic field due to a current carrying conductor. This will be done by following the same procedure employed for calculating the electric field intensity **E** at a given point for an arbitrary charge distribution. In the procedure the charge distribution was divided into charge elements dq. Coulomb's law was then used to calculate the field contribution dE due to each element at the point in question. The field **E** was then obtained by integrating the field contributions dE for the entire distribution.

A similar procedure will now be described for computing **B** at any point due to an arbitrary current distribution. The current distribution is first divided into *current elements* and, using the law of Biot and Savart (described below), the field contribution d**B** due to each current element at the point in question is calculated. The field at that point is then obtained by integrating the field contributions for the entire distribution.

Fig. 8.1 shows an arbitrary current distribution consisting of a current i in a curved wire. The figure also shows a typical current element; it is a length dl of the conductor carrying current i. Its direction is that of the tangent to the conductor (dashed line). A current element cannot exist as an isolated entity; a way must be provided to lead the current into the element at one end and out of it at the other. Nevertheless, the actual circuit can be thought of as made up of a large number of current elements placed end to end.

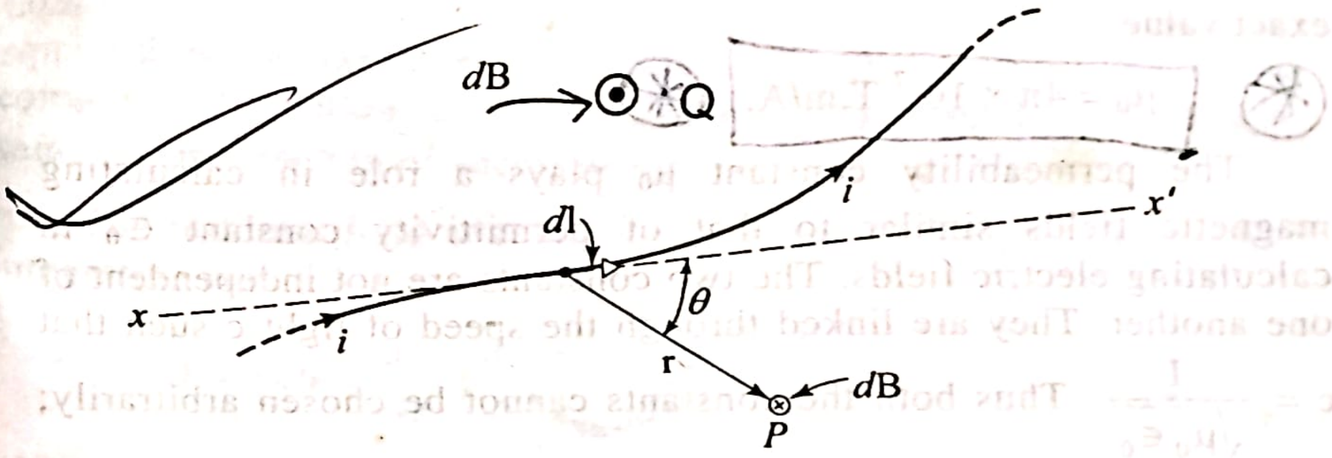


Fig. 8.1

From their investigations of the force produced by a current carrying conductor, Biot and Savart showed that the strength of the magnetic field dB at a point P due to the current element dl is

(i) proportional to the length dl of the element,

(ii) inversely proportional to the square of the distance r of the point from the element,

(iii) proportional to the sine of the angle θ between the direction of the element and the line joining the mid-point of the element to the point and

(iv) proportional to the strength i of the current

Thus

$$dB \propto \frac{i\, dl \sin\theta}{r^2}$$

or, $$dB = k_m \frac{i\, dl \sin\theta}{r^2} \qquad (8.1)$$

where k_m is a numerical constant. k_m in SI units is defined to have the exact value 10^{-7} tesla metre per ampere (Tm/A). However, as was the case in electrostatics, we find it convenient to write the constant in a different form:

$$k_m = \frac{\mu_0}{4\pi} = 10^{-7} \text{ T.m/A}$$

where the constant μ_0, called the *permeability constant*, has the exact value

$$\mu_0 = 4\pi \times 10^{-7} \text{ T.m/A.}$$

The permeability constant μ_0 plays a role in calculating magnetic fields similar to that of permittivity constant ϵ_0 in calculating electric fields. The two constants are not independent of one another. They are linked through the speed of light c such that $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$. Thus both the constants cannot be chosen arbitrarily; one of the constants can be set arbitrarily, but then the other is determined by the accepted value of c. Eqn. 8.1, therefore, reduces to

$$dB = \frac{\mu_0 i}{4\pi} \frac{dl \sin\theta}{r^2} \quad (8.2)$$

The direction of the field is found by grasping the wire with the right hand, the thumb pointing in the direction of **dl**. The curled fingers will then show the direction of **dB**; it is at right angles to the plane containing **r** and **dl**. At the pint P, **dB** points into the plane of the figure at right angles, as shown by a cross. Therefore, the direction of **dB** is that of the vector **r** × **dl**. At a point Q on the other side of the wire (in the same plane) **dB** points out of the plane and is shown by a dot.

It may be noted that eqn. 8.2, being an inverse square law that describes the magnetic induction due to a current element, may be viewed as the magnetic equivalent of Coulomb's law that describes the electric field due to a charge element

Eqn. 8.2, can be written as

$$dB = \frac{\mu_0 i}{4\pi} \frac{dl \sin\theta \times r}{r^2 \times r}$$

$$= \frac{\mu_0 i}{4\pi} \frac{dl \sin\theta r}{r^3}$$

$$\text{or,} \quad dB = \frac{\mu_0 i}{4\pi} \frac{dl \times r}{r^3} \tag{8.3}$$

Eqn. (8.3) is the vector form of Biot and Savart's law. It reduces to eqn. 8.2 when expressed in terms of magnitudes; it also gives complete information about the direction of **dB**, namely that it is the same as the direction of vector **dl × r.**

The resultant field at P due to the complete circuit is found by integrating eqn. 8.3. Or,

$$B = \int dB = \frac{\mu_0 i}{4\pi} \oint \frac{dl \times r}{r^3} \tag{8.4}$$

where the integral is a vector integral.

Applications of the Biot- Savart law

(i) Magnetic field due to a long straight wire :

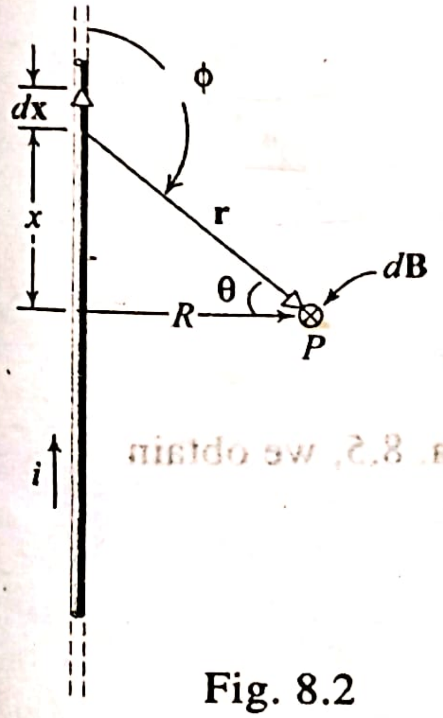


Fig. 8.2

Biot–Savart law can be illustrated by applying it to find the magnetic field B due to a current i in a long straight wire. Fig. 8.2 shows a typical current element dx. The magnitude of the cotribution dB of this element to the magnetic field at P is found by applying Biot–Savart law.

$$\boxed{dB = \frac{\mu_0 i}{4\pi} \frac{dx \sin\varphi}{r^2}}$$

where φ is the angle between the current element dx and the displacement vector r from the element to P. Since $\varphi = 90° + \theta$ where θ is the angle as shown in the figure

$$dB = \frac{\mu_0 i}{4\pi} \frac{dx \sin(90° + \theta)}{r^2}$$

$$= \frac{\mu_0 i}{4\pi} \frac{dx \cos\theta}{r^2} \qquad (8.5)$$

The direction of **dB** is that of vector **dx** × **r**, namely, perpendicular into the plane of the figure indicated by a cross sign. It may be noted that all such current elements of the wire give contributions in this same direction. So it needs only to compute the magnitude of the field.

In order to sum over all the current elements, we need to relate the variables θ, r and x. It turns out to be easiest to eliminate x and r in favour of θ. We have

$$x = R \tan\theta$$

Then $dx = R \sec^2\theta . d\theta$

$$= R(1 + \tan^2\theta)\, d\theta$$

$$= R\left(1 + \frac{x^2}{R^2}\right) d\theta$$

$$= R\left(\frac{R^2 + x^2}{R^2}\right) d\theta$$

$$= R\left(\frac{r^2}{R^2}\right) d\theta$$

$$= \frac{r^2}{R} . d\theta$$

Substituting this expression for dx into eqn. 8.5, we obtain

$$dB = \frac{\mu_0 i}{4\pi} \frac{r^2 d\theta}{R} \frac{\cos\theta}{r^2}.$$

$$= \frac{\mu_0 i}{4\pi} \frac{\cos\theta d\theta}{R}$$

Let us first calculate the contributions from the current elements below the point x = 0. We sum over these elements by integrating form $\theta = 0$, to $\theta = \theta_1$ where θ_1, is the angle between the line perpendicular to the wire and the line from P to the lower end of the wire as shown in the figure (Fig. 8.2)

We, therefore, have for this contribution,

$$B_1 = \int_0^{\theta_1} \frac{\mu_0 i}{4\pi} \frac{\cos\theta d\theta}{R} = \frac{\mu_0 i}{4\pi} \frac{\sin\theta_1}{R}$$

Similarly the contribution for the current elements above the point x = 0 is

$$B_2 = \frac{\mu_0 i}{4\pi} \frac{\sin\theta_2}{R}$$

Thus the total magnetic field due to the wire is

$$B = B_1 + B_2 = \frac{\mu_0 i}{4\pi R} (\sin\theta_1 + \sin\theta_2)$$

If the wire is very long, $\theta_1 = \theta_2 \simeq 90°$.

Therefore, the magnetic field produced by an infinitely long wire carrying a current i is

$$B = \frac{\mu_0 i}{4\pi R} (\sin 90° + \sin 90°)$$

$$= \frac{\mu_0 i 2}{4\pi R} = \frac{\mu_0 i}{2\pi R} \qquad (8.6)$$

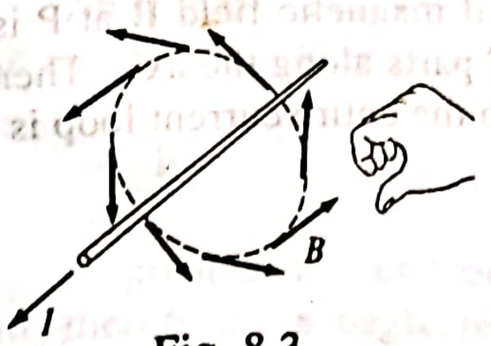


Fig. 8.3

The direction of **B** is such that the lines of **B** encircle the wire as shown in Fig. 8.3 which can be obtained by applying the right–hand rule.

(ii) Field due to a circular current loop:

Fig. 8.4 shows a circular loop of radius R carrying a current i. We would like to calculate the magnetic field **B** at a point P on the axis of the loop a distance z from the center of the loop.

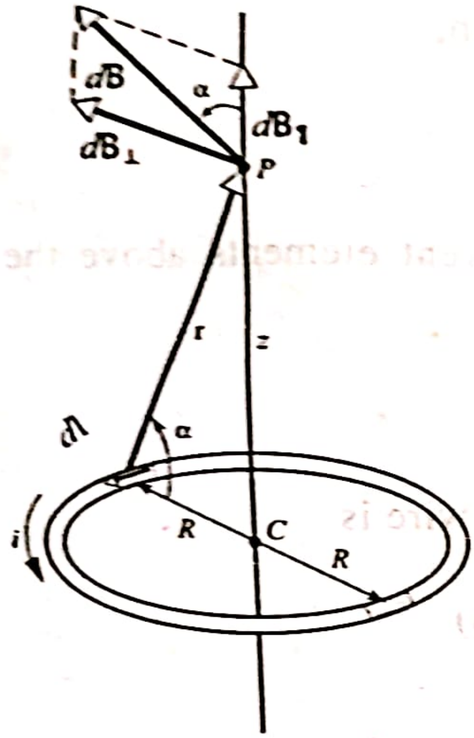


Fig. 8.4

Let us first consider a small element dl as shown in the figure. The magnitude of the magnetic field dB produced by this element at the point P is given by Biot-Savart law to be

$$dB = \frac{\mu_0 i}{4\pi}\,\frac{dl \sin 90^\circ}{r^2}$$

where the angle between the element idl and **r** is 90°. The field dB is perpendicular to the plane formed by idl and **r** and thus lies at right angle to r, as the figure shows.

dB may be resolved into two components, one, dB_{11} along the axis of the loop, and another $dB_\perp$, at right angles to the axis. For every element **dl** in the upper half of the circular current, there is an element at the opposite end (in the lower half) of a diameter, whose field at P will give a component dB_{11} along the axis. But the perpendicular component $dB_\perp$ will be equal and opposite to that due to the elément at the upper half. Thus the components perpendicular to the axis cancel and the sum of all $dB_\perp$ for the complete loop is zero. Only the components along the axis add and the total magnetic field **B** at P is simply the sum of all dB_{11} *i.e.*, the resolved parts along the axis. Then the total magnetic field produced at P due to the entire current loop is

$$B = \int dB = \int dB_{11}$$

Now $dB_{11} = dB \cos\alpha = \dfrac{\mu o i}{4\pi}\,\dfrac{dl \cos\alpha}{r^2}$

Hence

$$B = \int dB_{\parallel} = \int \frac{\mu_0 i}{4\pi} \frac{dl \cos\alpha}{r^2} \tag{8.7}$$

Fig. 8.4 shows that r and α are not independent of each other. Let us express each in terms of z, the distance of the point P from the center of the loop.

The relationships are

$$r = \sqrt{R^2 + z^2}$$

and $\cos\alpha = \dfrac{R}{r} = \dfrac{R}{\sqrt{R^2 + z^2}}$

Substituting these values in eqn. 8.6 we obtain

$$B = \int \frac{\mu_0 i dl}{4\pi (R^2 + z^2)} \frac{R}{\sqrt{R^2 + z^2}} \tag{8.8}$$

Now i, R and z have the same values for all current elements. Hence eqn. 8.8 can be written as

$$B = \frac{\mu_0 i R}{4\pi (R^2 + z^2)^{3/2}} \int dl$$

Now $\int dl$ is simply the circumference of the loop (= 2πR). Hence

$$B = \frac{\mu_0 i R . 2\pi R}{4\pi (R^2 + z^2)^{3/2}} = \frac{\mu_0 i R^2}{2 (R^2 + z^2)^{3/2}} \tag{8.9}$$

An important special case of this result is the magnetic field at the centre of the loop (z = 0). Eqn. then reduces to

$$B = \frac{\mu_0 i}{2R} \tag{8.10}$$

At great distances from the loop, z is much greater than R. R^2 can, therefore, be neglected compared to z^2 in the denominator of eqn. 8.9 Eqn. 8.9 then becomes

$$B = \frac{\mu_0 i R^2}{2z^3} \tag{8.11}$$

For a tightly wound coil of N identical circular loops, the total field is N times the value given by eqn. 8.11.

$$B = \frac{\mu_0 i N R^2}{2z^3} = \frac{\mu_0 i N \pi R^2}{2\pi z^3} = \frac{\mu_0 i N A}{2\pi z^3} \qquad (8.12)$$

where $A = \pi R^2$ is the area of a loop.

Remembering that $NiA = \mu$, the *magnetic dipole moment* of the loop, we can write

$$B = \frac{\mu_0}{2\pi}\frac{\mu}{z^3} \qquad (8.13)$$

8.2 Forces between two parallel current carrying conductors

Soon after Oersted's discovery that a magnetic needle is deflected by a current-carrying conductor, Ampere concluded that two such conductors would exert force of magnetic origin on one another.

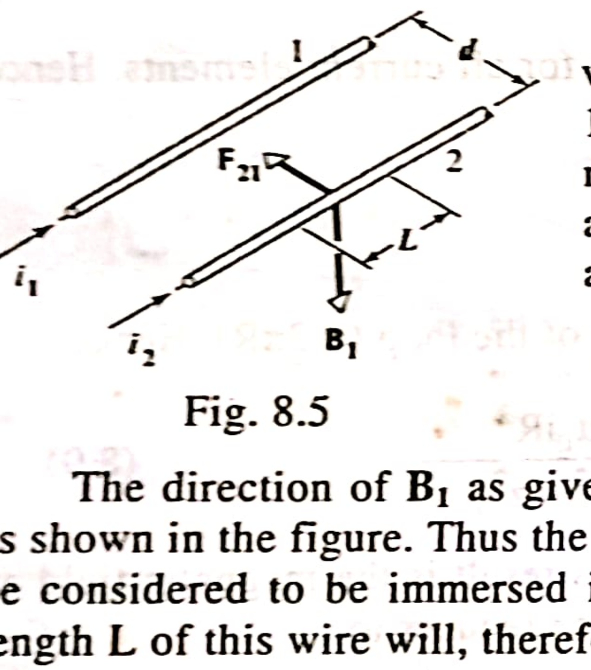


Fig. 8.5

Fig. 8.5 shows two long parallel wires separated by distance d. Wire 1, carrying a current i_1, produces a magnetic field B_1 whose magnitude at the site of second wire is, according to eqn. 8.5,

$$B_1 = \frac{\mu_0 i_1}{2\pi d}$$

The direction of $\mathbf{B_1}$ as given by right-hand rule, is downwards as shown in the figure. Thus the wire 2 which carries a current i_2 can be considered to be immersed in an external magnetic field $\mathbf{B_1}$. A length L of this wire will, therefore, experience a sideways magnetic force $F_{21} = i_2\, \mathbf{L} \times \mathbf{B_1}$ of magnitude

$$F_{21} = i_2 L B_1 \sin 90° = \frac{\mu_0 i_1 i_2 L}{2\pi d} \qquad (8.14)$$

where F_{21} stands for the force on wire 2 due to the magnetic field produced by wire 1. The vector rule for the cross product shows that $\mathbf{F_{21}}$ lies in the plane of the wires and points towards the first wire.

Instead of starting with 1, we could have started equally well with

wire 2 by first computing the magnetic field B_2 produced by wire 2 at the site of wire 1 and then computing the sideways fore F_{12} exerted on a length L of wire 1 by the field of wire 2. For parallel currents, this force on wire 1 would point towards wire 2. Thus the forces that the two wires exert on each other are equal in magnitude and opposite in direction; as they must be according to Newton's third law of action and reaction. The two wires attract each other.

If the currents in Fig. 8.5 were anti-parallel it would be that the forces on the wires would have opposite directions; the wires would repel one another. The general rule is:

Parallel currents attract, and anti-parallel currents repel

In a sense this rule is opposite to the rule for electric charges, in that like (parallel) currents attract, but like (same sign) charges repel. However, the magnetic interaction between two currents is similar in a sense to the electric interaction between two charges. One charge sets up an electric field, the other charge then interacts with the field at its particular location. Similarly the current in a wire sets up a magnetic field, and the current in the other wire then interacts with the field. The two phenomena may be represented as

$$\text{charge} \leftrightarrows \mathbf{E} \leftrightarrows \text{charge}$$

and

$$\text{current} \leftrightarrows \mathbf{B} \leftrightarrows \text{current.}$$

The Ampere

The force between two long parallel wires is used to define the *ampere*, the unit of current. If in eqn. 8.14, $i_1 = i_2 = 1$ ampere, d = 1 metre, L = 1 metre; $\mu_0 = 4\pi \times 10^{-7}$ Web/amp-m, then

$$\frac{F}{L} = \frac{\mu_0 i^2 L}{2\pi d L} = \frac{(4\pi \times 10^{-7}\,\text{Web/amp} - \text{m})\,(1\,\text{amp})^2}{(2\pi)\,(1\,\text{metre})}$$

$$= 2 \times 10^{-7}\ \text{nt/m}.$$

Thus *the ampere is the strength of that constant current which flowing through two parallel, straight and very long conductors of negligible cross-section and placed in vacuum one meter apart produces a force of 2×10^{-7} nt/m on unit length of each wire.*

The definition of ampere also gives us a method for measuring the value of μ_0. If $i_1 = i_2 = 1$ amp, $L = d = 1$ meter and $F = 2 \times 10^{-7}$ nt/m, then

$$\frac{\mu_0}{2\pi} = F = 2 \times 10^{-7} \text{ nt/m}$$

or, $\mu_0 = 4\pi \times 10^{-7}$ Web/amp-m.

The coulomb constant k $\left(= \dfrac{1}{4\pi\epsilon_0}\right)$ is 9×10^9 N–m^2/C^2 and the magnetic constant k_m ($= \mu_0/4\pi$) is 10^{-7} Web/amp-m. Hence the ratio of $\dfrac{k}{k_m} = \dfrac{1/4\pi\epsilon_0}{\mu_0/4\pi} = \dfrac{9\times10^9 \text{ N}-\text{m}^2/\text{C}^2}{10^{-7} \text{ Web/amp}-\text{m}} = 9 \times 10^{16}$ m^2/sec^2.

The ratio equals the square of the speed of light, a fact noted by Maxwell in 1860. He showed that according to the laws of electricity and magnetism an accelerated charge should radiate energy in the form of waves which travel with a speed

$$c = \sqrt{\frac{k}{k_m}} = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3 \times 10^8 \text{ m/sec.}$$

Since the speed turns out to be the same as the measured speed of light, Maxwell speculated that light is an electromagnetic wave produced by the acceleration of charges.

***Example 8.1** What is the magnetic field at a point 50 mm from a wire carrying a current of 3A?*

Soln.

Assuming a long wire, we have

$$B = \frac{\mu_0 i}{2\pi r} \qquad r = 50 \text{ mm} = 0.05\text{m}$$

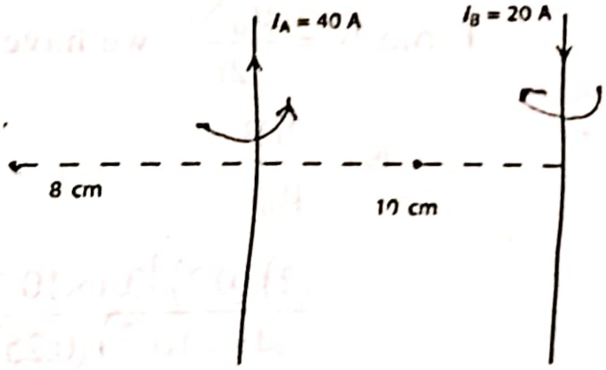
$$= \frac{(4\pi\times10^{-7} \text{ Web/amp}-\text{m})(3\text{A})}{(2\pi)(0.05 \text{ m})} = 1.20 \times 10^{-5}\text{T} = 0.12\text{G.}$$

***Example 8.2** Two long and fixed parallel wires, A and B are 10 cm apart in air and carry currents of 40 and 20 A, respectively in*

opposite directions. Determine the resultant flux density (a) on a line midway between the wires and parallel to them, and (b) on a line 8 cm form Wire A and 18 cm from wire B.

Soln.

(a) At the midpoint between in wires the fields both point into the page and hence reinforce.



$B = B_A + B_B$

$$= \frac{\mu_0 i_A}{2\pi r} + \frac{\mu_0 i_B}{2\pi r} \qquad r = 0.05 \text{ m}$$

$$= \frac{(4\pi\times10^{-7})(40+20)}{(2\pi)(0.05)}$$

$= 2.4 \times 10^{-4}$T.

(b) Here B_A points out of the page and B_B points into the page.

$$\therefore \quad B = \frac{(4\pi\times10^{-7})}{(2\pi)}\left[\frac{40}{0.08} - \frac{20}{0.18}\right] = 7.8 \times 10^{-5}\text{T}$$

out of the page.

***Example 8.3** A circular coil of radius 40mm consists of 250 turns of wire in which the current is 20 mA. What is the magnetic field in the centre of the coil?*

Soln.

At the centre of the coil,

$$B = \frac{\mu_0 N i}{2r}$$

N = 250
i = 20 mA = 20×10^{-3}A
r = 40 mm = 0.04m

$$= \frac{(4\pi\times10^{-7})(250)(20\times10^{-3})}{2\,(0.04)}$$

$= 0.785 \times 10^{-4}$T

= 0.785G.

Example 8.4 A coil of radius 200mm is to produce a field of 0.4G in its centre with a current of 0.25A. How many turns must there be in the coil?

Soln.

From $B = \frac{\mu_0 N i}{2r}$, we have

$$N = \frac{2rB}{\mu_0 i}$$

r = 200mm = 0.2m
B = 0.4G = 0.4 x 10^{-4}T
i = 0.25A

$$= \frac{(2)(0.2)(0.4\times10^{-4})}{(4\pi\times10^{-7})(0.25)}$$

= 51.

Example 8.5 A long, straight, non conducting string, painted with a charge at a density of 40 μC/m is pulled along its length at a speed to 300 m/s. What is the magnetic field at a normal distance of 5mm form the moving string?

Soln.

The moving string behaves just like a current with i = q x v = (40 x 10^{-6} C/m) (300 m/s) = 1.2 x 10^{-2}A.

$$\therefore \quad B = \frac{\mu_0 i}{2\pi r} = \frac{(4\pi\times10^{-7})(1.2\times10^{-2})}{(2\pi)(0.005)} = 4.8 \times 10^{-7}\text{T}.$$

Example 8.6 A long horizontal rigidly supported wire carries a current i_a of 100 amp. Directly above it and parallel to it is a fine wire that carries a current i_b of 20 amp and weighs 0.73 nt/metre. How far above the lower wire should this second wire be strung if we hope to support it by magnetic repulsion?

Soln.

For repulsion, the currents must point in the opposite directions. For equilibrium, the magnetic force per unit length must be equal to the weight per unit length of the wire and must be oppositely directed.

From $\frac{F}{l} = \frac{\mu_0 i_a i_b}{2\pi d}$ we have

$$d = \frac{\mu_0 i_a i_b}{2\pi(F/l)} = \frac{(4\pi\times10^{-7})(100)(20)}{(2\pi)(0.073)}$$

$$= 5.5 \times 10^{-3}\text{m} = 5.5 \text{ mm.}$$

8.3 Ampere's Law

Let us consider a hypothetical experiment of measuring the magnetic field **B** at various distances r from a long straight wire of circular cross-section and carrying a current i. This can be done by putting a small compass needle a distance r from the wire. Such a needle, which can be regarded as a small magnetic dipole, tends to line up with the external magnetic field, with its north pole pointing in the direction of **B**. It is obvious form Fig. 8.6 that **B** at the site of the dipole is tangent to a circle of radius r centered on the wire.

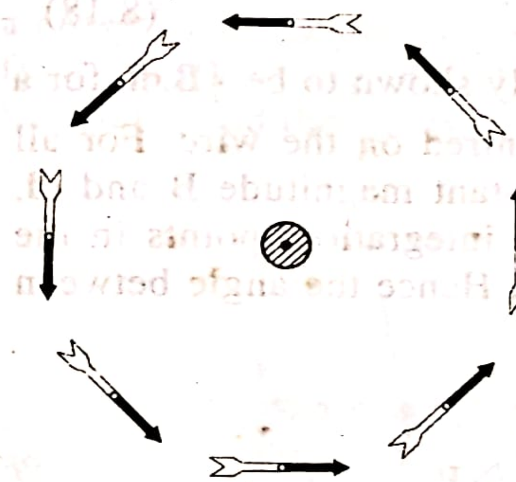


Fig. 8.6

If the direction of current is reversed, all the compass would reverse end-for-end. This experimental result leads to the *right-hand-rule* for finding the direction of **B** near a wire carrying a current i.

*Grasp the wire with the right hand, the thumb pointing in the direction of the current. The finger will curl around the wire in the direction of **B**.*

Let the dipole be now turned through an angle θ from its equilibrium position. To do this, we must exert an external torque just large enough to overcome the restoring torque τ that will act on the dipole. The relation between τ, θ and **B** is given by $\tau = \mu \times B$. In terms of magnitude this can be written as

$$\tau = \mu B \sin\theta \tag{8.15}$$

where μ is the magnitude of the magnetic moment of the dipole and θ is the angle between vectors **μ** and **B**. Although the actual value of **μ** for the compass needle may not be known, it may be taken to be a constant, independent of the position or orientation of the needle. Thus by measuring τ and θ in eqn. 8.15, the *relative* measures of B for various

distances r and for various currents i in the wire can be obtained. It has been found that the experimental results can be expressed by the relation.

$$B \propto \frac{i}{r} \tag{8.16}$$

The proportionality in eqn. 8.16 can be converted into an equality by inserting a proportionality constant. Instead of writing this simply as, say k, it is usually written in a more complex manner as $\mu_0/2\pi$, in which μ_0 is called the *permeability constant*. μ_0 has no relation with μ that appears in eqn. 8.15. Eqn. 8.16 then becomes

$$B = \frac{\mu_0}{2\pi}\frac{i}{r} \tag{8.17}$$

Eqn. 8.17 can be written in the form

$$(B)(2\pi r) = \mu_0 i \tag{8.18}$$

The left side of eqn. 8.18 can be easily shown to be $\oint \mathbf{B}.d\mathbf{l}$ for a path consisting of a circle of radius r centred on the wire. For all points on this circle **B** has the same constant magnitude B and dl, which is always tangential to the path of integration, points in the same direction as **B**, as shown in Fig. 8.7. Hence the angle between **B** and **dl** is zero. Thus

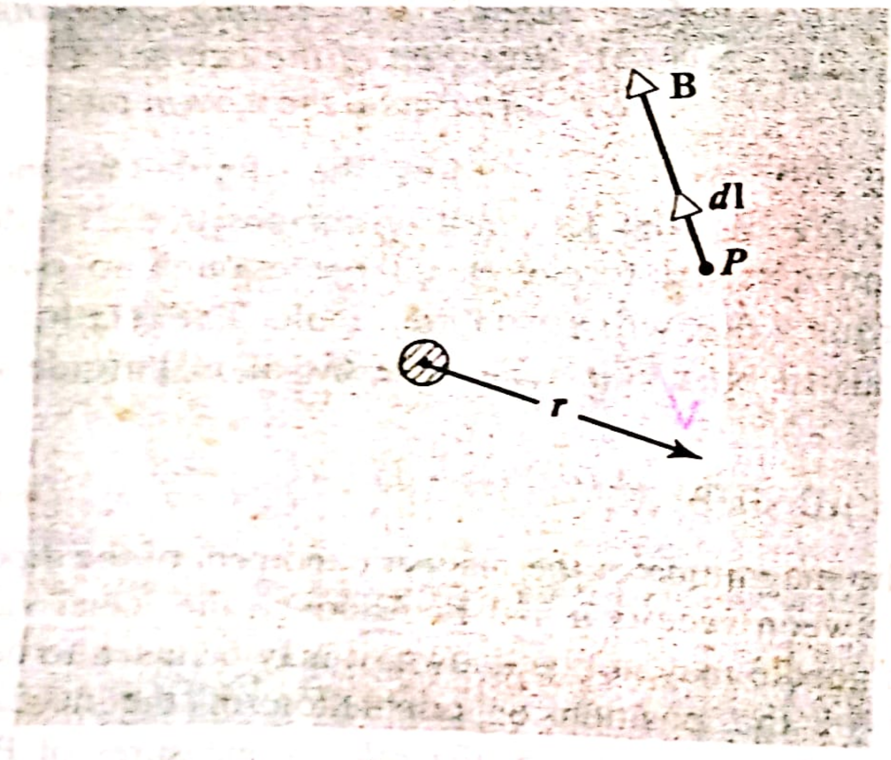


Fig. 8.7

$$\oint \mathbf{B}.d\mathbf{l} = \oint B\,dl \cos 0^\circ$$

$$= B\oint dl = (B)\,(2\pi r)$$

$\oint dl$ being simply the circumference of the circle.

In the special case of the magnetic field produced by a long, straight wire carrying a current i, the experimentally observed connection between the field and the current can be written as

$$\oint \mathbf{B}.d\mathbf{l} = \mu_0 i \tag{8.19}$$

Eqn. 8.19 is the *Ampere's law*. Although derived on the basis of the magnetic field produced by a long current currying conductor, a host of other experiments suggest that eqn. 8.19 is true in general for *any* magnetic field configuration, for *any* assembly of currents, and for *any* path of integration.

In applying Ampere's law, the general procedure is to construct a *closed linear path* – the so called *Amperean loop*, in the magnetic field shown in Fig. 8.8. The path is then divided into elements of length dl, and the quantity **B**.dl is evaluated for each element. Now

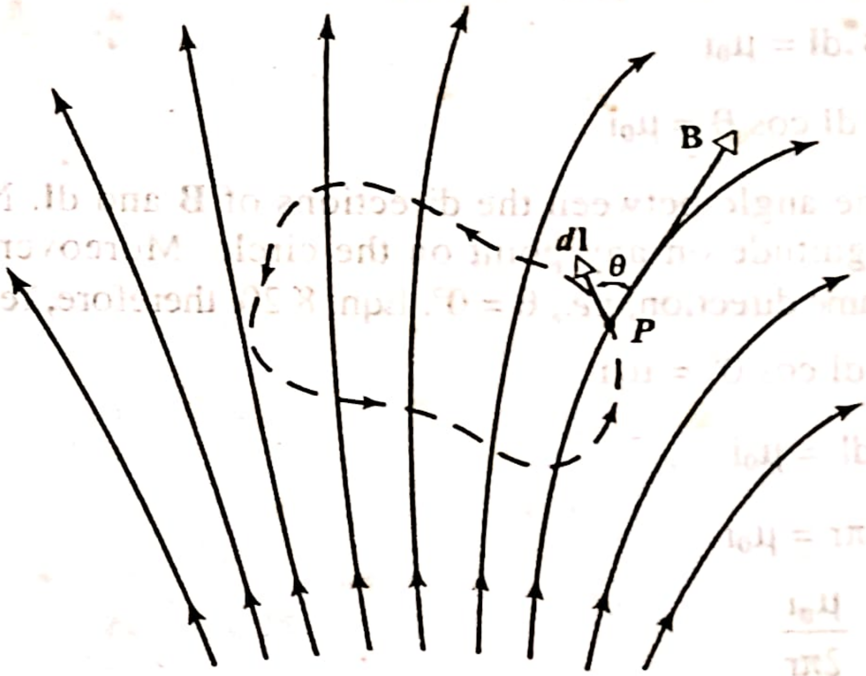


Fig. 8.8

B.**dl** has the magnitude Bdlcosθ, θ being the angle between the directions of **B** and **dl**. Hence **B**.**dl** can be interpreted as the product of dl and the component of **B** (= B cos θ) parallel to **dl**. The integral is the sum of the quantities **B**.**dl** for all path elements in the complete loop. The term i on the right hand side of eqn. 8.19 is the *net* current that passes through the area bounded by the closed path.

The permeability constant in Ampere's law has an assigned value of

$$\mu_0 = 4\pi \times 10^{-7} \text{ Weber/amp-metre.}$$

The permeability constant μ_0 and the permittivity constant ϵ_0 occur in electromagnetic formulae when MKS system of units is used.

Application of Ampere's Law

(i) B near a long wire

A long straight wire is carrying a current i. We want to find the magnetic field **B** at a point P a distance r from the wire.

The lines of magnetic induction for a long straight wire carrying a current i are concentric circles centered on the wire. Hence the point P may be regarded to be on a circular loop of radius r surrounding the wire. Applying Ampere's law, we have

$$\oint \mathbf{B}.\mathbf{dl} = \mu_0 i$$

$$\text{or,} \quad \oint B\, dl \cos\theta = \mu_0 i \qquad (8.20)$$

where θ in the angle between the directions of **B** and **dl**. Now **B** has the same magnitude on any point on the circle. Moreover, **B** and **dl** point in the same direction, *i.e.*, θ = 0°. Eqn. 8.20, therefore, reduces to

$$\oint B\, dl \cos 0^\circ = \mu_0 i$$

$$\text{or,} \quad B\oint dl = \mu_0 i$$

$$\text{or,} \quad B.2\pi r = \mu_0 i$$

$$\text{or,} \quad B = \frac{\mu_0 i}{2\pi r} \qquad (8.21)$$

the integral $\oint dl$ being simply the circumference of the circle of radius r.

Eqn. 8.21 is similar to that obtained by applying Biot-Savart law. But here, the same expression has been derived with much more ease and elegance.

(ii) The magnetic field of a solenoid

A *solenoid* is a long wire wound in a closed packed helix and carrying a current i. It is assumed that the helix is very long compared with its diameter. What is the nature of the magnetic field B that is set up by the helix?

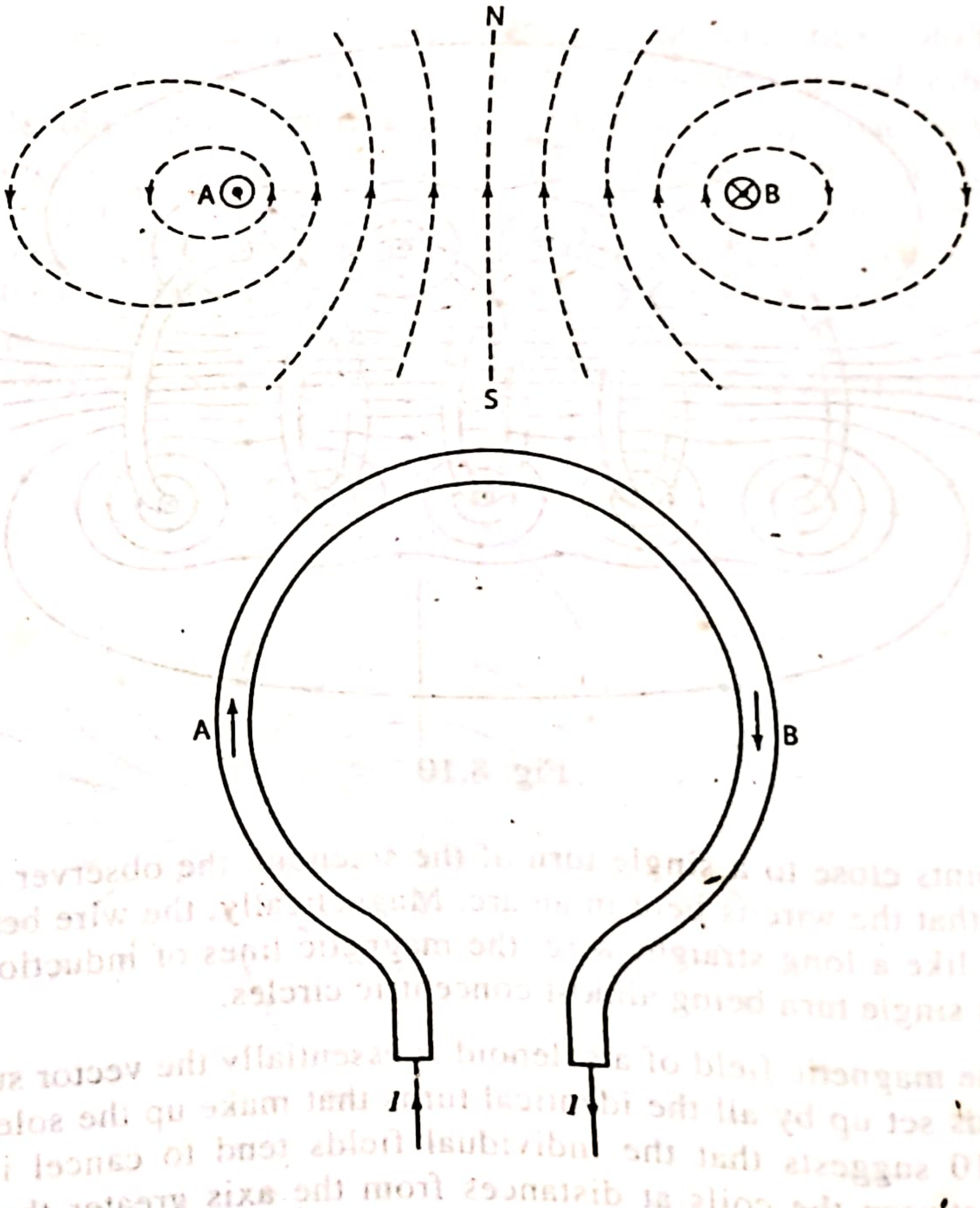


Fig. 8.9

Let us consider the case of a single turn coil (Fig. 8.9). The flux distribution along the section AB is shown in the upper of the diagram. Since the current flows *upwards* at point A, then the flux path at that point is anticlockwise. At point B the direction of the current is downwards, so that the flux path there is clockwise. The net result as shown in the figure is that a N–pole is formed at one side of the coil while a S-pole is formed on the other side.

The flux pattern associated with a *solenoid* or a multi-turn coil can be deduced from the foregoing discussion. Fig. 8.10 shows, for the sake of illustration, only a section of a *stretched out* solenoid.

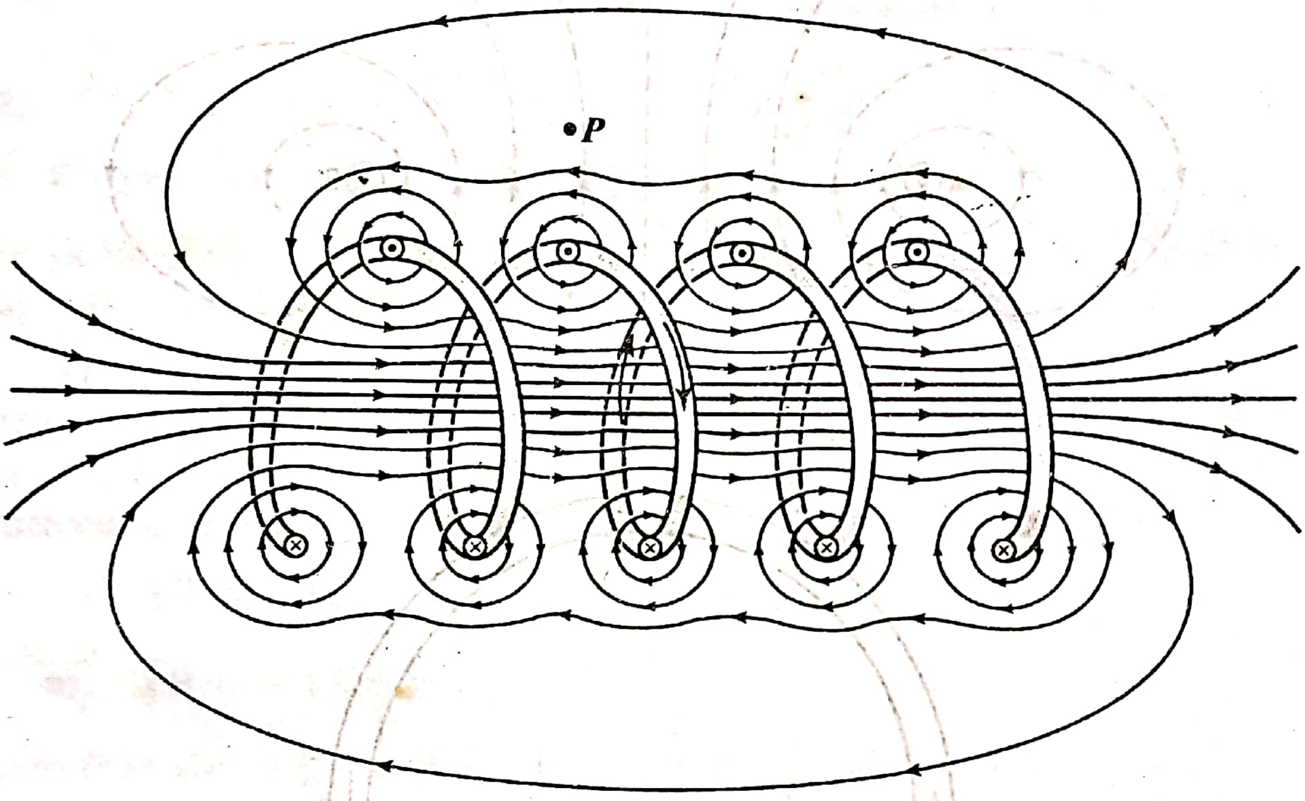


Fig. 8.10

For points close to a single turn of the solenoid, the observer is not aware that the wire is bent in an arc. Magnetically, the wire behaves almost like a long straight wire, the magnetic lines of induction (**B**) for this single turn being almost concentric circles.

The magnetic field of a solenoid is essentially the vector sum of the fields set up by all the identical turns that make up the solenoid. Fig. 8.10 suggests that the individual fields tend to cancel in the space between the coils at distances from the axis greater than the

coil radius. But in the space between the coils near the axis, the fields of the individual coils add, and reasonably far from the wires, **B** is parallel to the solenoid axis. In the limiting case of a tightly packed square wires, the solenoid essentially becomes a cylindrical current sheet, and the requirements of symmetry then make it rigorously true that **B** is parallel to the axis of the solenoid. This will be assumed to be the case in the following discussion.

Referring to Fig. 8.10, the field set up by the upper part of the solenoid turns near points such as P (marked ⊙, because the current is out of the page) points to the left. The field set up by the lower part of the solenoid turns (marked ⊗, because the current is into the page) points to the right. These two fields tend to cancel each other. As the solenoid becomes more and more ideal, *i.e.*, it approaches the configuration of an infinitely long cylindrical current sheet, the field **B** at outside points approaches zero. For a practical solenoid, whose length is much greater then its diameter, it is a good approximation to take the external field to be zero, specially if we consider the external points near the central region of the solenoid, *i.e.*, away form the ends. Fig. 8.11, shows the lines of magnetic induction for a

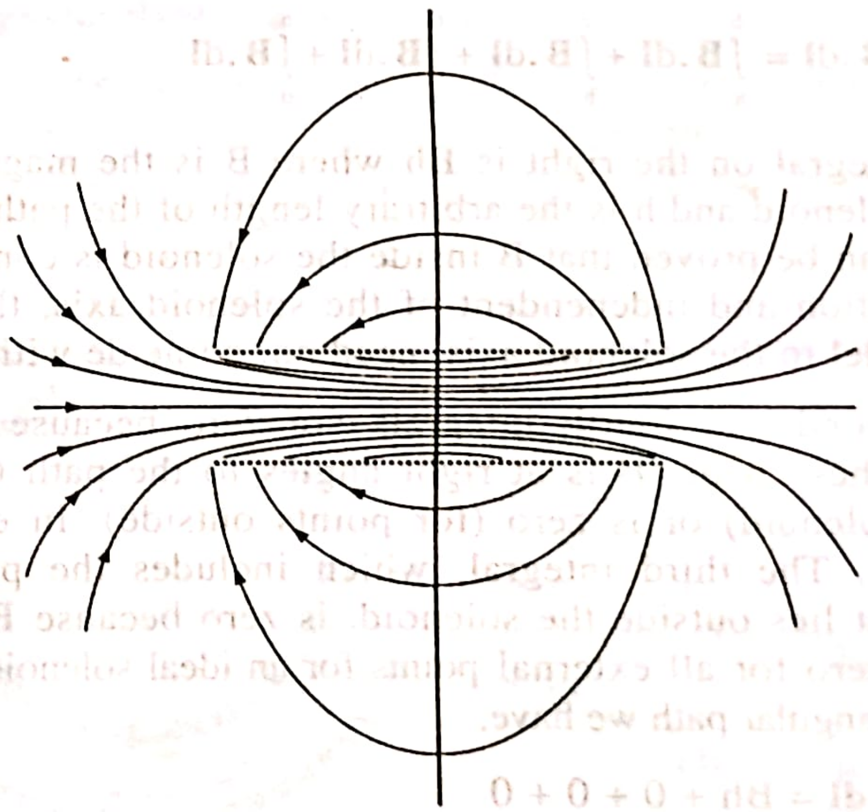


Fig. 8.11

real solenoid which is far from ideal in that the length is not much greater than the diameter. Even here it is obvious that the external field in the central plane is much weaker than the internal field as indicated by the spacing of the magnetic lines of induction.

Let us apply Ampere's law

$$\oint \mathbf{B}.d\mathbf{l} = \mu_0 i$$

to the rectangular path abcd in the ideal solenoid of Fig. 8.12. The integral $\oint \mathbf{B}.d\mathbf{l}$ can be written as the sum of four integrals, one for each path segment:

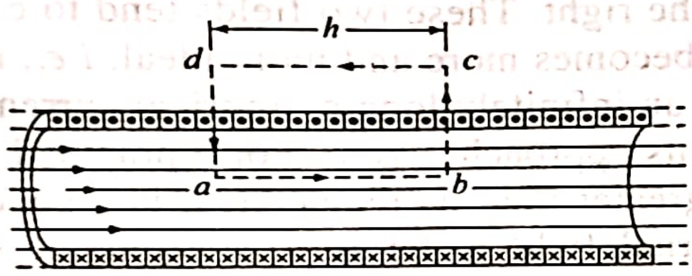


Fig. 8.12

$$\oint \mathbf{B}.d\mathbf{l} = \int_a^b \mathbf{B}.d\mathbf{l} + \int_b^c \mathbf{B}.d\mathbf{l} + \int_c^d \mathbf{B}.d\mathbf{l} + \int_d^a \mathbf{B}.d\mathbf{l}$$

The first integral on the right is Bh where B is the magnetic field inside the solenoid and h is the arbitrary length of the path from a to b. Since it can be proved that B inside the solenoid is constant over its cross–section and independent of the solenoid axis, the path ab though parallel to the solenoid axis, need not coincide with it.

The second and fourth integrals are zero because for every element of these paths **B** is at right angles to the path (for points inside the solenoid) or is zero (for points outside). In either case **B.dl** is zero. The third integral, which includes the part of the rectangle that lies outside the solenoid, is zero because **B** has been taken to be zero for all external points for an ideal solenoid. Thus for the entire rectangular path we have,

$$\oint \mathbf{B}.d\mathbf{l} = Bh + 0 + 0 + 0$$

$$= Bh = \mu_0 i$$

The net current that passes through the Amperean loop is not the same as the current i_0 in the solenoid because the windings pass through the loop more than once. If n is the number of turns per unit length of the solenoid, then the current passing through the loop is

$$i = i_0 n h$$

Ampere's law then becomes

$$Bh = \mu_0 i_0 n h$$

or, $$B = \mu_0 i_0 n \qquad (8.22)$$

Eqn. 8.22 shows that the magnetic field inside a solenoid depends only on the current i_0 and the number of turns per unit length n.

Although eqn. 8.22 was derived for an ideal solenoid, it holds quite well for actual solenoids for internal points near the centre of the solenoid. It suggests that B does not depend on the diameter or the length of the solenoid and that B is constant over the solenoid cross section. A solenoid is a practical way to set up a known uniform magnetic field for experimentation, just as a parallel-plate capacitor is a practical way to set up a known uniform electric field.

(iii) *Magnetic field due to a toroid*

Fig. 8.14 shows a toroid which may be considered to be a solenoid bent into the shape of a doughnut. We shall find the magnetic field at interior points by applying Ampere's law and certain consideration of symmetry.

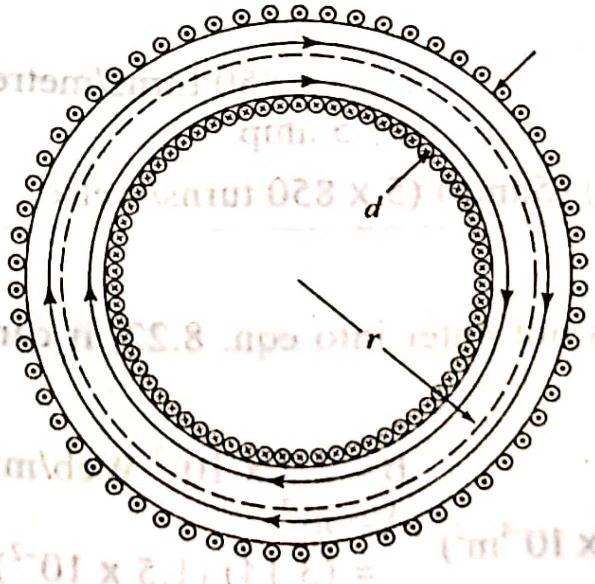


Fig. 8.13

From symmetry, the lines of **B** form concentric circles inside the toroid, as shown in the figure. Let us choose a concentric circle of radius r as an Amperean loop and traverse it in the clockwise direction. Applying Ampere's law we get

$$\oint \mathbf{B} . d\mathbf{l} = \mu_0 i$$

or, $$(B)(2\pi r) = \mu_0 i_0 N$$

where i_0 is the current in the windings and N is the total number of turns. This gives

$$B = \frac{\mu_0 i_0 N}{2\pi r} \qquad (8.23)$$

In contrast to the solenoid, B is not constant over the cross-section of the toroid. It can be shown by applying Ampere's law that B = 0 for points outside an ideal toroid.

The denominator in eqn. 8.23, $2\pi r$, is the central circumference of the toroid, and $N/2\pi r$ is just n, the number of turns per unit length. With this substitution, eqn. 8.23 becomes

$$B = \mu_0 i_0 n$$

the equation for the magnetic field in the central region of a solenoid.

The direction of the magnetic field within a toroid (or a solenoid) follows from the right-hand rule: *curl the fingers of your right hand in the direction of the current; the extended thumb then points in the direction of the magnetic field.*

***Example 8.7** A solenoid is 1.0 metre long and 3.0 cm in mean diameter. It has five layers of windings of 850 turns each and carries current of 5.0 amp. (a) What is B at its centre? (b) What is the magnetic flux φ_B for a cross-section of the solenoid at its centre?*

Soln.

$n = 5 \times 850$ turns/metre

$i_0 = 5$ amp

(a) $B = \mu_0 i_0 n$

$= (4\pi \times 10^{-7}$ Web/amp-m) (5amp) (5×850 turns/metre)

$= 2.7 \times 10^{-2}$ Web/metre^2

As the diameter of the winding does not enter into eqn. 8.22, it can be used to find B.

$B = 2.7 \times 10^{-2}$ Web/m^2

$A = \pi r^2$

$= (3.14)\ (1.5 \times 10^{-2})$

$= 7.1 \times 10^{-4}$m^2.

(b) $\varphi_B = \int \mathbf{B}.d\mathbf{s} = BA$

$= (2.7 \times 10^{-2}$ Web/m^2) (7.1×10^{-4}m^2)

$= 1.9 \times 10^{-5}$Wb.

***Example 8.8** An air-core solenoid with 2000 loops on it is 600 mm long and has diameter of 20 mm. If a current of 5A is sent through it, what will be the flux density within?*

Soln.

$$B = \mu_0 i_0 n \qquad N = \frac{2000}{0.6m}$$

$$= (4\pi \times 10^{-7} \text{ Wb/amp-m})(5 \text{ amp})\left(\frac{2000}{0.6m}\right)$$

$$= 0.021 \text{ T.}$$

For a long (30:1) solenoid such as this, the field will be 21mT everywhere inside, except right near the ends.

***Example 8.9** A solenoid 0.5m long has 2000 turns. The magnetic induction near the centre of the solenoid is 0.08T. What is the current in the solenoid?*

Soln.

For a long solenoid,

$$B = \mu_0 i_0 n$$

$$\text{or,} \quad i_0 = \frac{B}{\mu_0 n} \qquad n = \frac{2000}{0.5m} = 4000 \text{ m}^{-1}$$

$$= \frac{0.08T}{\left(4\pi \times 10^{-7} \text{ Web/amp} - \text{m}\right)\left(4000 \text{ m}^{-1}\right)}$$

$$= 16 \text{ A.}$$

EXERCISES

1. State Biot-Savart law.

   Apply Biot-Savart law to find the magnetic filed due to a long straight wire carrying a current.

2. State and explain Biot-Savart law.

   Find the value of B due to a current loop of radius R carrying a current i at a distance x on the axis from the centre of the loop. What will be the value of B at the centre of the loop?

3. Obtain an expression for the force between two parallel current carrying conductors. Under what condition the force becomes attractive or repulsive?

4. Derive an expression for the force between two parallel current carrying conductors and hence define an ampere.

5. State and explain Ampere's law.

   Apply Ampere's law to obtain an expression for the force per unit length between two parallel current carrying conductors. When will the force be attractive and when repulsive? Hence obtain a definition for the ampere.

6. State Ampere's law.

   What is a solenoid? Find the magnetic field **B** that is set up in a solenoid having n number of turns per unit length and carrying a current i_0.

7. Apply Ampere's law to find the magnetic field due to long current carrying conductor.

8. What is a toroid? Apply Ampere's law to find the magnetic field at the centre of a toroid.

9. A solenoid has a mean diameter of 0.03m and length 2 metres. It has four layers of 1000 turns each. Calculate the flux density at its centre when a current of 2.5 amperes flows through it. Also calculate the flux at the centre. **[6.28×10^{-3} Web/m^2, 1.23×10^{-5} Wb]**

10. Calculate the flux density at the centre of a coil if there are 100 turns wound on a solenoid 25cm long and carrying a current of 3 ampere. **[1.508×10^{-3} Wb/m^2]**

11. Two long parallel wires are 10 cm apart in air and carry currents of 6A and 4A respectively. Compute the force of attraction on each meter length of wire if the currents are (a) in the same direction and (b) in opposite direction. **[4.8 x 10^{-5}N attraction, 4.8 x 10^{-5}N repulsion]**

12. A charge of 4×10^{-4}C moves at 15 revolutions per second in a circle of radius 10 cm. What is the magnetic flux density at the centre of the circle? **[1.88 x 10^{-10}T]**

13. A circular coil of 40 turns of wire and negligible section has diameter 32 cm. What current must exist in the coil to produce a flux density of 3×10^{-4} T at its centre? **[1.9 A]**

14. A wire carrying a current of 100 amperes is bent into the form of a circle of radius 5.08 cm. Calculate (a) flux density at the centre of the coil (b) flux density perpendicular to the plane of the coil at a distance of 19 cm from the coil. **[12.37 x 10^{-4} Wb/m^2, 7.33 x 10^{-5} Wb/m^2]**

15. A flat circular coil having 40 loops of wire on it has a diameter of 320 mm. What current must flow in its wires to produce a flux density of 300 μWb/m^2 at its centre? **[1.9 A]**

16. A solenoid has a length L = 1.23m and an inner diameter d = 3.55cm. It has five layers of winding of 850 turns each and carries a current i_0 = 5.57A. What is B at its centre? **[24.2 x 10^{-3}T]**

17. Five very long, straight, insulated wires are closely bound together to form a small cable. Currents carried by the wires are I_1 = 20A, I_2 = – 6A, I_3 = 12A, I_4 = – 7A, I_5 = 18A. (negative currents are opposite in direction to the positive). Find B at a distance of 10 cm from the cable. **[74 μT]**

18. A coil of radius 2 cm is concentric with a coil of radius 7cm. Each coil has 100 turns. With a current of 5A in the larger coil, find the currents needed in the smaller coil to give the following values for the total magnetic field at the centre: (a) 9.0mT, (b) 2.0 mT and (c) zero. In each case determine whether the direction of the current in the smaller coil is the same as the current in the larger coil or opposite. **[1.44A and has the same direction, 0.793A and has the opposite direction, 1.43A and has the opposite direction.]**

CHAPTER - IX

ELECTROMAGNETIC INDUCTION

The discovery by Oersted in 1820 that a magnetic field is always associated with an electric current led the physicists of the day to ask: *'If a steady electric current produces a steady magnetic field, should not a steady magnetic field produce a steady electric current'*. All efforts to find such a current were unsuccessful, although both Ampere and Arago observed the phenomenon but failed to recognize it. However, Joseph Henry, an American physicist, observed in 1830 that current flows in a coil around the iron bridge (or keeper) across the pole of his electromagnet when current was started or stopped in the coils of the magnet itself. Faraday was quick to recognize that the question posed above should be rephrased as follows: *'if moving electric fields, that is those of moving charges that constitute the current, produce magnetic fields, should we not expect that moving magnetic fields would produce electric currents'*. In 1831 Faraday was able to detect the production of current in a coil of wire wound on an iron core when, and only when, another current was started or stopped in another coil on the same core. Faraday's prompt publication and clarification of the phenomenon gave him priority of discovery over Henry who did not publish his observations until Faraday's publication. This phenomenon is called *electromagnetic induction* and the currents thus generated *induced currents*.

9.1 Faraday's experiments

Consider a coil of wire whose terminals are connected to a galvanometer as shown in Fig. 9.1. Normally one should not expect any deflection in the galvanometer because there seems to be no source of electromotive force in the circuit. However, if, now a permanent bar magnet is pushed toward the coil, with its north pole facing the coil, a remarkable thing happens. *While the magnet is moving*, the galvanometer deflects, showing that current has been set up in the coil. It the magnet is held stationary with respect to the coil, *i.e.*, there is no relative motion between the magnet and the coil, the galvanometer

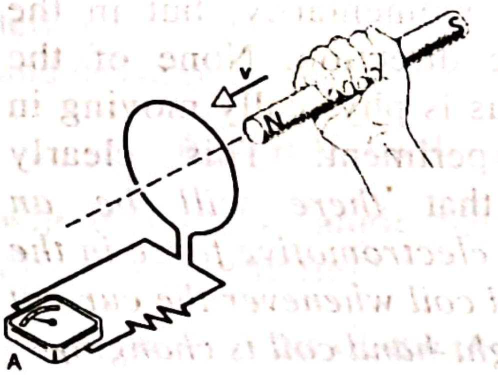


Fig. 9.1

does not deflect. If the magnet is moved *away* form the coil, the galvanometer again deflects, but in the opposite direction which means that direction of the current in the coil is now in the opposite direction. If the south pole of the magnet is used instead of the north pole, similar effects are observed in the coil, but now the deflections are reversed. The faster the magnet is moved toward the coil the greater is the deflection of the coil. But this deflection (and hence the current) lasts only for a shorter time. The presence of the current in the coil implies the existence of an electromotive force in the coil. Further experimentation shows that this *induced electromotive force* is present only when there is a *relative motion* of the magnet and the coil. *It makes no difference if the magnet is moved toward the coil or the coil toward the magnet.*

If the magnet in Fig. 9.1 is replaced by a current carrying solenoid, as in Fig. 9.2, then the movement of the solenoid with respect to the coil also produces an induced electromotive force in the coil.

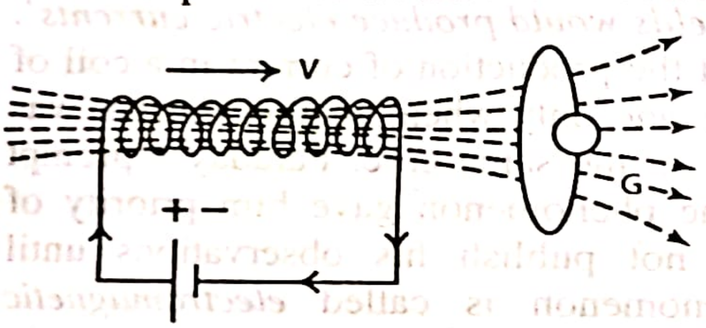


Fig. 9.2

In a third experiment performed by Faraday, the apparatus of Fig. 9.3 is used. Two coils are placed close together, face to face, but at rest with respect to each other. One coil contains a battery and a switch S, while the other a galvanometer G. When the switch S is closed, thus setting up a steady current in the right-hand coil, the galvanometer G in the left-hand coil deflects momentarily; when the switch is opened, thus disturbing the steady current in the right-hand coil, the galvanometer again deflects momentarily, but in the opposite direction. None of the apparatus is physically moving in this experiment. This clearly shows that *there will be an induced electromotive force in the left-hand coil whenever the current in the right-hand coil is changing.*

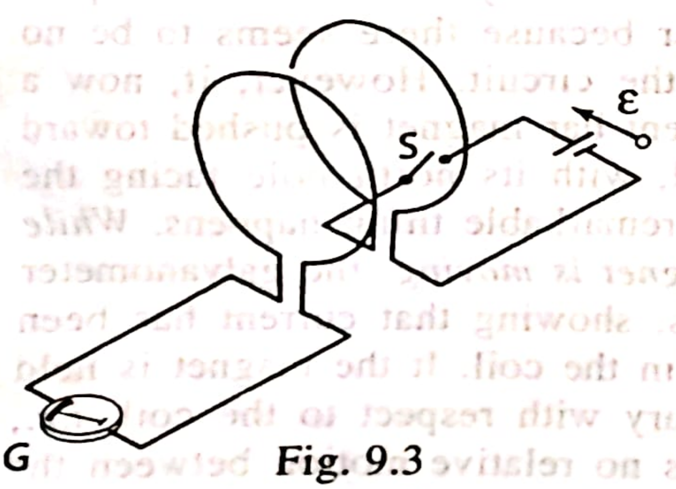


Fig. 9.3

The common feature of these experiments is *motion* or *change*. It is the moving magnet or the *changing* current that is responsible for the induced electromotive forces. It is the rate at which the current is changing and not the size of the current, that is significant.

9.2 Laws of electromagnetic induction

Faraday had the insight to perceive that it is the change in the lines of magnetic induction linked with the bar magnet or the current carrying coil in the preceding experiments which is the important common factor. Imagine that there are lines of magnetic field coming from the bar magnet of Fig. 9.1 or the current carrying coil of Fig. 9.2. Some of these lines pass through the left-hand coil, connected to the galvanometer, in both figures. As the magnet is moved towards this coil as in Fig. 9.1 or the switch S is opened or closed in Fig. 9.3, the number of lines of magnetic field passing through the left hand coil changes. From his experiments Faraday concluded that, *it is the change in the number of field lines passing through a circuit loop that induces the electromotive force in the loop*. Specifically, it is the *rate of change* in the number of field lines passing through the loop that determines the induced emf.

Let us now introduce the term *magnetic flux* ϕ_B . Like electric flux, magnetic flux may be considered to be the measure of the number of field lines passing through a surface. If dA be an clement of area at a point where the magnetic field strength is **B** then the flux passing through dA is

$$\phi_B = \int \mathbf{B} . d\mathbf{A}$$

where the integration is carried out over the entire surface through which we wish to calculate the flux, for example, the surface enclosed by the left-hand loop in Fig. 9.1 If the magnetic field has a constant magnitude and direction over planar area A, the flux can be written as,

$$\phi_B = BA \cos \theta \qquad (9.1)$$

where θ is the angle between the normal to the surface and the direction of the field.

[The SI unit of magnetic flux is tesla.metre^2 which is given the name of *weber* (abbreviation Wb). That is

$$1\ \text{Wb} = 1\ \text{tesla.metre}^2$$

$$\text{or,}\quad 1\ \text{tesla} = 1\ \text{Wb/metre}^2$$

Weber used to be the unit of magnetic field before tesla was adopted as the unit in SI units]

In terms of the magnetic flux, the laws of electromagnetic induction are given below.

Faraday's law

An induced *e.m.f.* is established in a circuit whenever the magnetic flux linking that circuit is changed.

Neumann's law

The magnitude of the induced *e.m.f.* is proportional to the rate of change of the magnetic flux linking the circuit.

Lenz's law

The induced *e.m.f.* acts to circulate a current in a direction that oppose the change in flux which caused the induced *e.m.f.*

All these three laws may be summarized in a single law which states that

the electromotive force induced in a circuit is equal to the negative of the time rate of change of the magnetic flux through the circuit.

The above statement is usually referred to as *Faraday's law of electromagnetic induction.* In mathematical terms, Faraday's law can be written as

$$\varepsilon = -\frac{d\varphi_B}{dt} \qquad (9.2)$$

where ε is the electromotive force induced in the circuit. If the rate of change of flux is in units of webers per second, then the *e.m.f.* has the unit of volts.

Eqn. 9.2 is one of the basic laws of electricity. The minus sign in

eqn. 9.2 is very important, because it tells us the direction of the induced *e.m.f.* The significance of the minus sign will be considered in more detail in Art. 9.3.

If the coil consists of N turns, then eqn. 9.2 can be applied to each turn; an *e.m.f.* appears in every turn and the total *e.m.f.* in the circuit is the sum of these *e.m.fs*, just as in the case of batteries connected in series. If the coil is so tightly wound such that each turn occupies the same region in space, the flux through each turn will be same (as are the cases with solenoids and toroids). The total induced *e.m.f.* will then be given by.

$$\varepsilon = -N\frac{d\varphi_B}{dt} = -\frac{d(N\varphi_B)}{dt}$$

where $N\varphi_B$ measures the so called *flux-linkage* in the circuit

There are many ways by which the flux linked with a loop may be changed. These are: moving a magnet relative to the loop (as in Fig. 9.1), changing the current in a nearby circuit (as in Fig. 9.3) moving the loop in a non-uniform field, rotating the loop in a fixed magnetic filed such that the angle θ in eqn. 9.1 changes (as in generator), changing the size or shape of the loop. In each of these cases, an *e.m.f.* is induced in the circuit. It may be noted that even though eqn. 9.2 is known as Faraday's law, it was not written in that form by Faraday, who was untrained in mathematics. The mathematical form of the law is referred to as Neumann's Law. In fact, Faraday's three volumes of published work on electromagnetism, a landmark achievement in the development of physics and chemistry, contains not a single equation.

9.3 Lenz's Law

The negative sign in Faraday's law indicates that direction of the induced e.m.f. is such as to tend to oppose the change that produces it. Thus if an attempt is made to increase the flux through a circuit, the induced e.m.f. tends to cause currents in such a direction as to decease the flux. If an attempt is made to thrust the north pole of a magnet towards a current carrying loop, the face of the loop toward the magnet must become a north pole (Fig. 9.4). The two poles - one of the current loop and one of the magnet - repel each

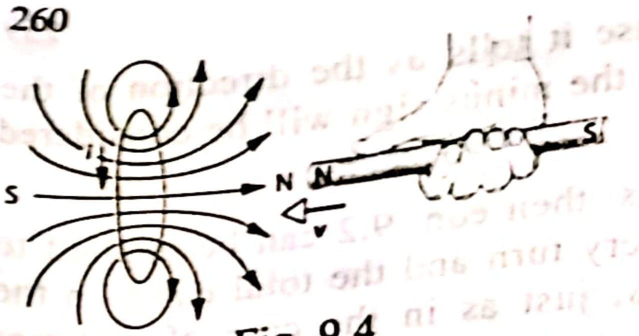


Fig. 9.4

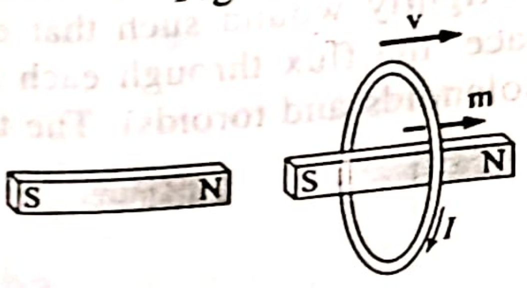


Fig. 9.5

other. On the other hand, if the coil is pulled away from the magnet, the induced current opposes the *pull* by creating a *south* pole on the face of the loop near the magnet (Fig. 9.5). All these phenomena are covered

by the law proposed in 1834 by Heinrich Friedrich Lenz which states that

the direction of the induced emf is such as to oppose the cause producing it.

In the statement of the Lenz's law the kind of change that causes the induced e.m.f. was not specified. This was left purposefully vague in order to allow for a variety of interpretations all of which are valid and give the same result for the direction of the induced e.m.f and current. For example, if we consider the case mentioned above, the movement of the magnet towards the coil tend to increase the flux through the coil. The induced e.m.f. must oppose this increase of flux through the coil. Hence the face of the loop near the magnet becomes a north pole so as to oppose the movement of the magnet. If the induced e.m.f. in Fig. 9.4 were in the *opposite* direction to that shown, as magnet moves towards the loop, the face of the loop near the magnet would be a south pole, which would *pull* the bar magnet toward the loop. We would only need to push the magnet slightly to start the process and then the action would be *self-perpetuating*. The magnet would accelerate toward the loop, increasing the kinetic energy all the time. At the same time internal energy (Joule heating) would appear in the loop at a rate that would increase with time. This would indeed be a *something-for-nothing* situation! Needless to say it does not occur. The induced field must oppose the change that gives rise to it. The direction of the induced e.m.f must be such so that the face of the loop near the magnet becomes a north pole so as to oppose the movement of the magnet.

9.4 Motional Electromotive force

Consider a rectangular loop of wire of width l, one end of which is in a uniform field **B** pointing at right angles to the plane of the loop as shown in Fig. 9.6. The magnetic field **B** may be produced, for example, in the gap of a large electromagnet. The dashed lines show the assumed limits of the magnetic field. The loop is pulled to the right at a constant speed v.

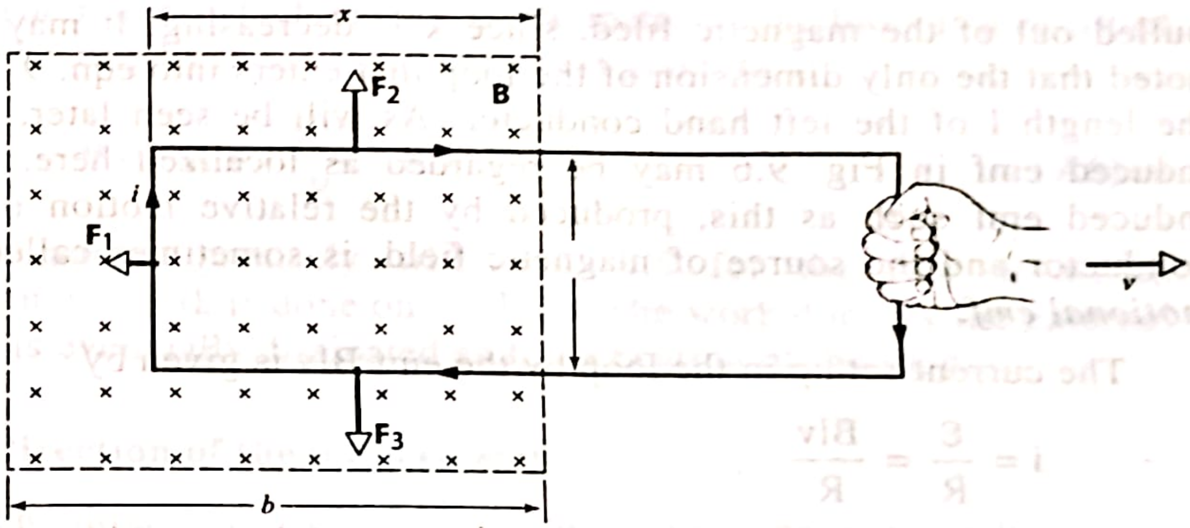


Fig. 9.6

The situation as depicted in Fig. 9.6 does not differ in any essential detail from that of Fig. 9.5. In each case a conducting loop and a magnet are in relative motion; difference between the two situations is that, although the situation depicted in Fig. 9.5 is easy to understand qualitatively, it does not lend itself to quantitative calculations. The situation of Fig. 9.6, however, permits easier calculations.

The external agent (the hand in Fig. 9.6) pulls the loop to the right at constant speed v by exerting a force F. We would like to calculate the rate at which it does work on the loop, and to compare the result with the rate at which the induced current in the loop produces internal energy.

The flux φ_B enclosed by the loop in Fig. 9.6 is

$$\phi_B = B.A = Blx$$

where lx is the area of that part of the loop in which B is not zero. The electromotive force $\mathcal{E}$ produced is found from Faraday's law. Or

$$\mathcal{E} = -\frac{d\varphi_B}{dt} = -\frac{d}{dt}(Blx)$$

$$= -Bl\frac{dx}{dt} = +Blv \qquad (9.3)$$

where $-\frac{dx}{dt}$ has been set equal to the speed v at which the loop is pulled out of the magnetic filed, since x is decreasing. It may be noted that the only dimension of the loop that enters into eqn. 9.3 is the length l of the left hand conductor. As will be seen later, the induced emf in Fig. 9.6 may be regarded as localized here. An induced emf such as this, produced by the relative motion of a conductor and the source of magnetic field, is sometimes called a *motional emf.*

The current set up in the loop by the emf Blv is given by

$$i = \frac{\mathcal{E}}{R} = \frac{Blv}{R}$$

where R is the resistance of the loop. From Lenz's law this current and hence the emf $\mathcal{E}$ must be clockwise in Fig. 9.6; it opposes the *change* (the decrease of ϕ_B) by setting up a field that is parallel to the external field within the loop.

The current in the loop gives rise to three magnetic forces $\mathbf{F_1}$, $\mathbf{F_2}$ and $\mathbf{F_3}$ that act on the three sides of the loop in accordance with eqn. 7.13 of chapter vii.

$$\mathbf{F} = i\,\mathbf{l} \times \mathbf{B}$$

Since $\mathbf{F_2}$ and $\mathbf{F_3}$ are equal and opposite, they cancel each other; $\mathbf{F_1}$ which is the force that opposes the effort to move the loop is given in magnitude by

$$F = i\,l\,B\sin\theta$$

$$= i\,l\,B\sin 90^\circ \qquad (\theta = 90^\circ)$$

$$= i\,l\,B$$

$$= \frac{Blv}{R}.lB = \frac{B^2l^2v}{R} \qquad (9.4)$$

The agent that pulls the loop must exert a force F equal in magnitude

to F_1 so that the loop may move at constant speed. The agent must therefore do work at the steady rate of

Work done per second,

$$P = F_1 v = \frac{B^2 l^2 v}{R} . v = \frac{B^2 l^2 v^2}{R} \qquad (9.5)$$

The rate at which energy is dissipated in the loop as a result of Joule heating by the induced current is given by

$$P = i^2 R = \left(\frac{Blv}{R}\right)^2 R = \frac{B^2 l^2 v^2}{R} \qquad (9.6)$$

which agrees precisely with eqn. 9.5 for the rate at which mechanical work is done on the loop. The work done by the external agent is eventually dissipated as Joule heating of the loop.

9.5 Direction of the induced emf

The direction in which the induced emf acts in a conductor can be obtained by applying Fleming's right-hand rule as follows:

"If the thumb, first finger and second finger of the right hand are held so that they point in directions that are mutually perpendicular (Fig. 9.7) then if the *F*irst finger points in the direction of the magnetic *F*lux, and the thu*M*b points in the direction of the *M*otion of the conductor, then the s*E*cond finger points in the direction of the induced *E*.m.f."

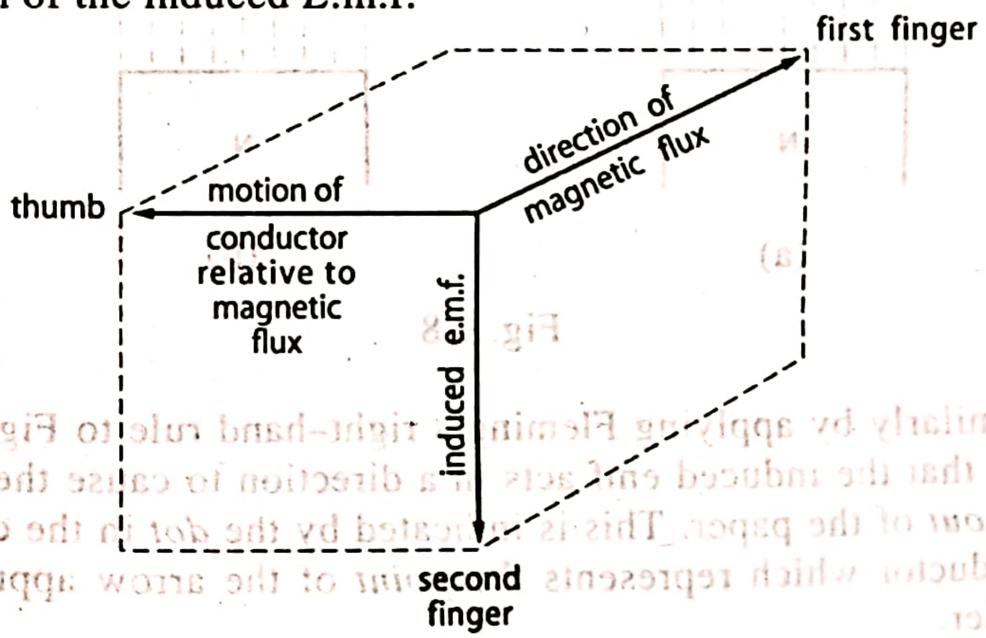


Fig. 9.7

That is

*F*irst finger	- direction of magnetic *F*lux
thu*M*b	- Motion the conductior *relative* to the flux
s*E*cond finger	- direction of *E*.m.f.

The application of Fleming's right-hand rule can be illustrated by considering the electro-mechanical systems of Fig. 9.8.

In Fig. 9.8(a) the conductor moves to the left relative to the magnetic field, and the direction of the magnetic flux is from the N–pole to the S–pole. Applying Fleming's right–hand rule we find that the direction of the induced emf would be such as to cause a current to flow *into* the paper. Since current flow is represented by an arrow, we indicate that the current is flowing away from the reader (*i.e.*, into the paper) by showing the *tail* or *crossed feathers* of the arrow on the conductor.

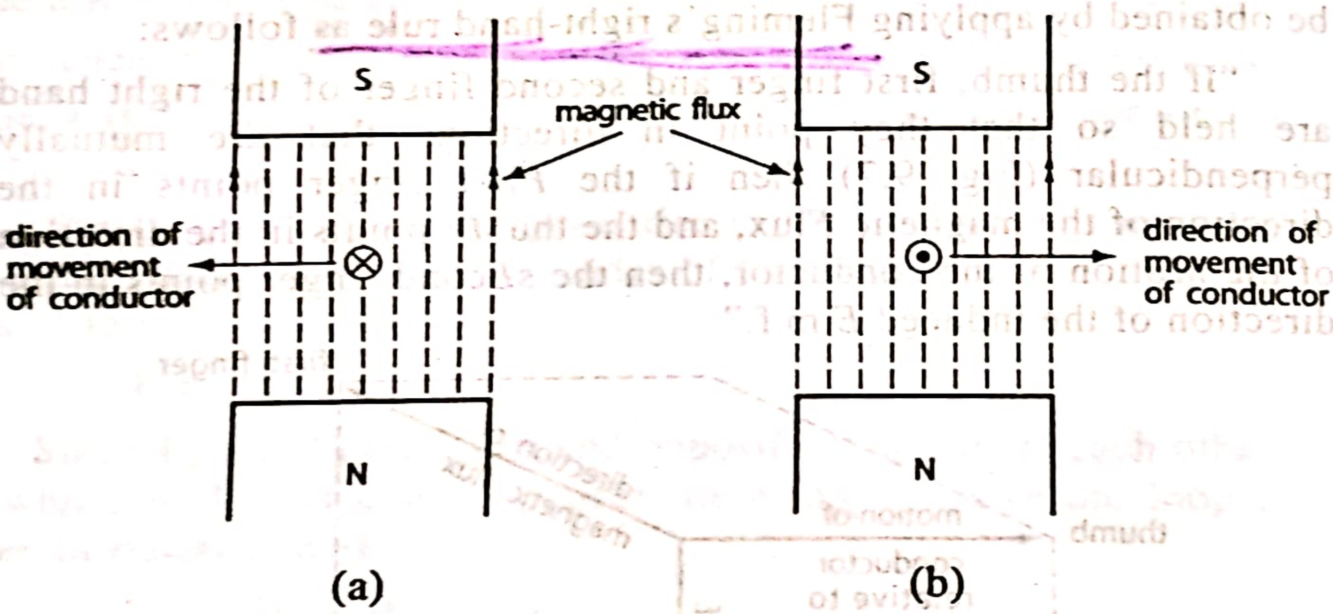


Fig. 9.8

Similarly by applying Fleming's right–hand rule to Fig. 9.8(b), we find that the induced emf acts in a direction to cause the current to flow *out* of the paper. This is indicated by the *dot* in the centre of the conductor which represents the *point* of the arrow approaching the reader.

Example 9.1 The flux linking an air–cored coil of 500 turns changes from 30μWb to 60μWb in 2ms. Calculate the value of the emf induced in the coil.

Soln.

$$d\varphi = (60 - 30)\ \mu Wb = 30 \times 10^{-6}\ Wb$$

$$dt = 2ms = 2 \times 10^{-3} s$$

$$\therefore \quad e = N\frac{d\varphi}{dt} = \frac{(500)\left(3\times10^{-6}\right)}{2\times10^{-3}} = 7.5\ V.$$

Example 9.2 A small search coil with an area of 125 mm² has 50 turns of very fine wire. This coil is placed between the pole pieces of a small magnet and then suddenly jerked out. If the average induced emf is 0.07V when the coil is pulled to a field-free region in 60ms, what is the magnetic intensity between the poles? What was the original flux through each turn?

Soln.

Let

φ = total flux linking the coil.

and φ_0 = flux through a single loop;

$$\therefore \quad \Phi = N\Phi_0$$

$$\text{Now } \mathcal{E} = \frac{\Delta\Phi}{\Delta t}; \quad \text{or} \quad 0.07 = \frac{(50)\left(125\times10^{-6}\right)B}{0.06}$$

where $\phi_0 = B.A$ and $A = 125\ mm^2 = 125 \times 10^{-6} m^2$.

Since the final flux is zero $\Delta\varphi = N\Phi_0$

and $\Delta t = 60ms = 0.06s$.

$$\text{or,} \quad B = \frac{42\times10^{-4}}{625\times10^{-5}} = 0.672\ T$$

$$\therefore \quad \phi_0 = BA = (0.672\ Wb/m^2)(125 \times 10^{-6} m^2)$$

$$= 84 \times 10^{-6}\ Wb = 84\ \mu Wb.$$

Example 9.3 The perpendicular component of the external magnetic field through a 10-turn coil of radius 50mm increases from 0 to 18T in 3s. If the resistance of the coil is 2 Ω, what is the magnitude of the induced current? What is the direction of the current?

Soln.

The initial flux Φ_1 is zero and the final flux

$$\Phi_2 = NBA \qquad N = 10;\ A = \pi r^2 = \pi(50 \times 10^{-3}m)^2$$

$$= (10)(18T)(25\pi \times 10^{-4}m^2) \qquad = 25\pi \times 10^{-4}m^2$$

$$= 1.41\ Wb. \qquad B = 18T.$$

So $\Delta\Phi = \Phi_2 - \Phi_1 = 1.41$ Wb and $\Delta t = 3s$.

The induced emf is

$$\mathcal{E} = -\frac{\Delta\Phi}{\Delta t} = -\frac{1.41\ Wb}{3\ s} = -0.47V.$$

The minus sign indicates that the emf would cause a current that creates a field opposing the change in B.

The magnitude of the induced current is

$$I = \frac{\mathcal{E}}{R} = \frac{0.47V}{2} = 0.235\ A.$$

Example 9.4 A 130-turn close-packed coil of diameter 2.1 cm is placed at the centre of a long solenoid which has 220 turns/cm and carries a current i = 1.5A; its diameter is 3.2 cm. The current in the solenoid is increased from zero to 1.5A at a steady rate over a period of 0.16s. What is the absolute value of the induced emf that appears in the central coil while the current in the solenoid is being changed?

Soln.

The absolute value of the final flux through each turn of the coil at the centre is

$\Phi_0 = BA \cos\theta = BA \quad (\theta = 0)$

The initial flux linked with this coil when i = 0 is zero.

The magnetic field B at the centre of the solenoid is given by

$$B = \mu_0 in$$

$$= (4\pi \times 10^{-7}\ \text{T.m/A})\ (1.5\ \text{A}) \times (220\ \text{turns/cm})\ (100\ \text{cm/m})$$

$$= 4.15 \times 10^{-2}\text{T}.$$

The area of a turn of the centre coil is given by

$$A = \pi r^2 = (\pi)\left(\frac{2.1}{2}\times 10^{-2}\,\text{m}\right)^2$$

$$= 3.46 \times 10^{-4}\text{m}^2.$$

The absolute value of the final flux through each turn of the coil is then

$$\Phi_B = BA = (4.15 \times 10^{-2}\text{T})\ (3.46 \times 10^{-4}\ \text{m}^2)$$

$$= 1.44 \times 10^{-5}\ \text{Wb} = 14.4\ \text{mWb}.$$

The magnitude of the induced emf is

$$\mathcal{E} = N\frac{\Delta\Phi_B}{\Delta t}$$

$N = 130$

$\Delta\Phi_B = 14.4\ \mu\text{Wb} - 0 = 14.4 \times 10^{-6}\ \text{Wb}.$

$\Delta t = 0.16\text{s}.$

$$= (130)\ (14.4 \times 10^{-6}\ \text{Wb})\ (0.16\text{s})$$

$$= 1.2 \times 10^{-2}\text{V} = 12\ \text{mV}.$$

Example 9.5 *A conductor of active length 0.3m moves perpendicular to a magnetic field at a velocity of 50 m/s. Calculate the average value of the emf induced in the conductor if the magnetic flux density is 0.5T.*

Soln.

$$\mathcal{E} = Blv = (0.5)\ (0.3)\ (50) = 7.5\ \text{V}.$$

***Example 9.6** The average emf induced in a conductor of length 0.6m is 0.5V. The conductor moves at an angle to a magnetic field of flux density 0.3T at a velocity of 10 m/s. Calculate the value of the angle between the direction of movement of the conductor and the direction of the magnetic field.*

Soln.

$$\varepsilon = Blv \sin\theta \qquad \varepsilon = 0.5V;\ B = 0.3\ T$$

$$l = 0.6m;\ v = 10\ m/s$$

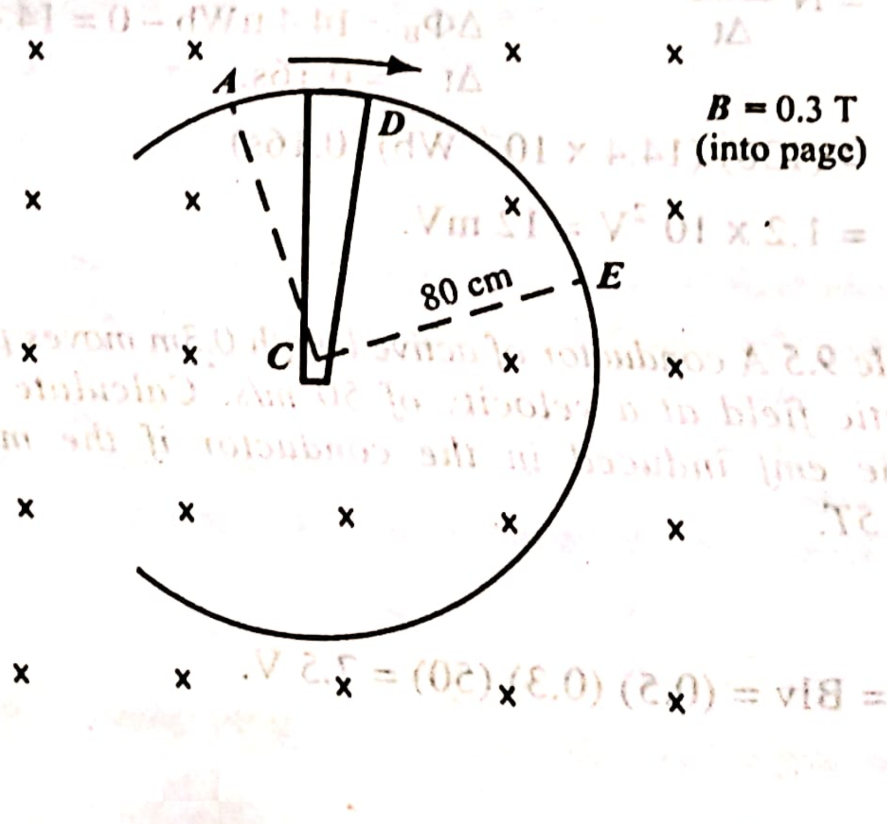
$$\text{or,}\quad \sin\theta = \frac{\varepsilon}{Blv}$$

$$= \frac{0.5}{(0.3)(0.6)(10)}$$

$$= 0.2778$$

$$\therefore\quad \theta = 16^\circ\ 08'.$$

***Example 9.7** The rod shown in the figure below rotates about point C as pivot with the constant frequency 5 rev/s. Find the potential difference between its two ends, 80 cm apart, because of the magnetic field, B = 0.3T.*



Soln.

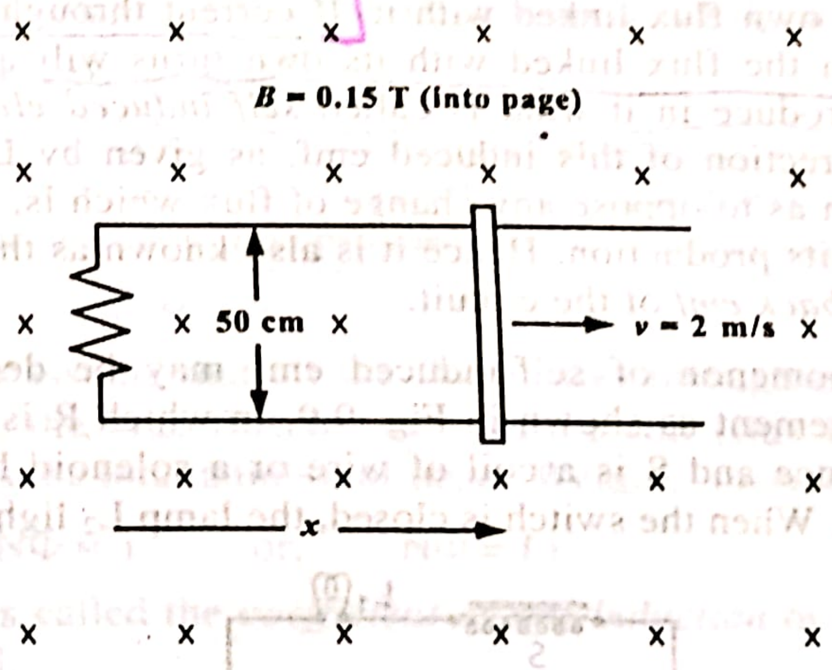
Let us consider a fictitious loop CADC. As time goes on its area will increase and so will the flux through it. The induced emf in this loop will be the potential difference between the two ends of the rod.

$$\mathcal{E} = N\frac{\Delta\phi}{\Delta t} = (1)\frac{B\Delta A}{\Delta t}$$

The area changes form zero to that of a full circle πr^2 in one -fifth of a second. Therefore,

$$\mathcal{E} = B\frac{\Delta A}{\Delta t} = B\frac{\pi r^2}{0.20s} = (0.3T)\frac{\pi(0.8m)^2}{0.20s} = 3.0\ V.$$

Example 9.8 As shown in the figure below, a metal rod makes contact with a partial circuit and completes the circuit. The circuit area is perpendicular to a magnetic field with B = 0.15T. If the resistance of the total circuit is 3 Ω, how large a force is needed to move the rod as indicated with a constant speed of 2 m/s?



Soln.

The induced emf in the rod will cause a current to flow counter clockwise in the circuit. Because of this current in the rod, it experiences a force to the left due to the magnetic field. In order to

pull the rod to the right with constant speed, this force must be balanced by the puller.

$$\mathcal{E} = Blv = (0.15T)\ (0.50m)\ (2\ m/s) = 0.15V.$$

$$I = \frac{\mathcal{E}}{R} = \frac{0.15\ V}{3\ \Omega} = 0.050\ A.$$

The mechanical force that is needed to move the rod is given by

$$F = ilB \sin 90° = \frac{B^2 l^2 v}{R}$$

$$ilB \sin 90° = (0.050\ A)\ (0.50m)\ (0.15T)\ (1) = 3.75 \times 10^{-3} N$$

Also, $$\frac{B^2 l^2 v}{R} = \frac{(.015T)^2\ (0.50m)^2\ (2m/s)}{3\ \Omega} = 3.75 \times 10^{-3} N.$$

9.6 Self-induced electromotive force

This is the electromotive force induced in a coil due to the change of its own flux linked with it. If current through the coil is changed, then the flux linked with its own turns will also change which will produce in it what is called *self induced electromotive force*. The direction of this induced emf, as given by Lenz's law, would be such as to oppose any change of flux which is, in fact, the very cause of its production. Hence it is also known as the *opposing* or *counter* or *back emf* of the circuit.

The phenomenon of self-induced emf may be demonstrated with an arrangement as shown in Fig. 9.9, in which R is a pure (or ohmic) resistance and S is a coil of wire or a solenoid having soft iron in its core. When the switch is closed, the lamp L_2 lights up first,

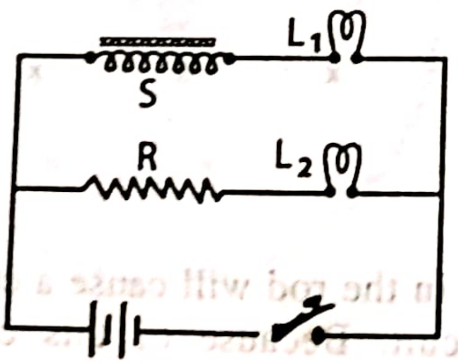


Fig. 9.9

L_1 afterwards. This is because when the key is closed, the current in the resistance R instantaneously attains its maximum value; as a result the lamp L_2 glows immediately. This is not the case with the coil S. Even without the presence of any other coil nearby, when a current is started in this circuit, the magnetic lines of induction so produced cause a change in induction from zero to maximum through the circuit *itself,* which gives rise to induced or back emf opposing the applied emf tending to drive the current. Because of this, the current in the circuit does not attain its steady maximum value instantaneously, and it may take time to do so, which may vary from a fraction of a second to several minutes. The lighting up of the lamp L_1 after the lamp L_2 demonstrates this slow growth of current in the inductive circuit.

Similarly at break an induced emf is set up, which tends to maintain the flow of current, and it may take some time before the current in the coil circuit is actually stopped. Thus the glow in L_2 is stopped earlier than in L_1.

This effect in the inductive coil is called *self-induction*. The coil, therefore possesses a property called *self-induction* or *electro-magnetic inertia*, the latter name being given as the mechanical inertia of a body, which tends to oppose the motion and stoppage of a body, is analogous to it.

Self-inductance

When a current i passes through a coil it will produce a flux Φ, which in turn, threading through the coil will, itself, induce an emf. The flux produced will be proportional to strength of the current.

Thus $\Phi \propto i$

If we assume that all the flux passes through each turn of the coil itself, the total flux linked with the coil (flux linkage) is NΦ, where N is the total number of turns. We may, therefore, write

$$N\Phi \propto i \qquad \text{or,} \qquad N\Phi = Li$$

where L is called the *coefficient of self-induction* or *self-inductance* of the coil.

The emf induced in the coil itself is given by

$$\mathcal{E} = -\frac{d}{dt}(N\Phi) = -L\frac{di}{dt} \tag{9.7}$$

This induced emf will produce an induced current which will be superposed on the main current.

Considering only the magnitudes, eqn. 9.7 can be written as

$$\mathcal{E} = L\frac{di}{dt} \tag{9.8}$$

$$\text{or,} \quad L = \frac{\mathcal{E}}{di/dt} = \frac{\text{induced emf}}{\text{rate of change of current}}$$

If $di/dt = 1$, $L = \mathcal{E}$

Thus the coefficient of self-induction or self-inductance *is the amount of emf induced between the ends of the coil when the rate of change of current in it is unity.*

Eqn. 9.8 shows that the SI unit of inductance is volt. second/ampere. This combination of unit has been given the special name of the *henry* (abbreviation H) so that

1 henry = 1 volt. second/ampere

This unit has been named after Joseph Henry, an American physicist and a contemporary of Faraday.

Thus a circuit has a self inductance of one henry if an emf of one volt is induced in the circuit when the current in the circuit changes at the rate of one ampere per second.

The relationship between the sign of $\mathcal{E}$ and the sign of di/dt can be obtained by applying Lenz's law. Let a steady current i be established in the solenoid (perhaps by a battery not shown in the figure) shown in Fig. 9.10.

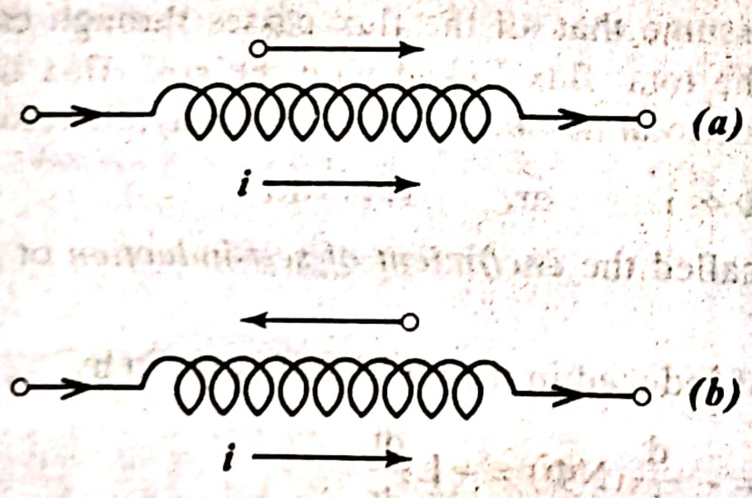


Fig. 9.10

Suppose the emf of the battery is suddenly increased; the current i at once starts to increase. According to Lenz's law this increase in current must be opposed by the self-inductance of the coil. Thus the emf induced in the coil due to this sudden increase in current sends a current in the direction opposite to i. Similarly when the current is suddenly decreased by decreasing the emf of the battery, Lenz's law requires that the emf induced in the coil oppose this decrease in current. Hence the emf induced sends a current in the same direction as i.

Calculation of inductance

The inductance can be calculated by a technique which is based on Faraday's law. First the magnetic field B is determined for the geometry of a particular inductor. This enables the magnetic flux ϕ_B through each turn of the coil to be obtained. Assuming that the flux has the same value for each of the N turns of the coil, the total flux through the coil is the product $N\Phi_B$, which is referred to as the flux linkage of the inductor. The emf induced is then found from Faraday's law.

$$\mathcal{E} = -\frac{d}{dt}(N\Phi_B)$$

The emf induced is also given by the relation

$$\mathcal{E} = -L\frac{di}{dt}$$

Comparing the two equations, we obtain

$$L\frac{di}{dt} = \frac{d}{dt}N\Phi_B \tag{9.9}$$

Integrating with respect to time we find

$$Li = N\Phi_B$$

$$\text{or,}\quad L = \frac{N\Phi_B}{i} \tag{9.10}$$

If the initial and final fluxes are Φ_1 and Φ_2 respectively and the corresponding currents are i_1 and i_2 then eqn. 9.10 can be written as

$$L = \frac{N(\Phi_2 - \Phi_1)}{i_2 - i_1} \quad (9.11)$$

Eqn. 9.11, which is based upon Faraday's law, permits the inductance to be found directly from the number of flux linkage. It may be noted that, since Φ_B is proportional to the current i, the ratio in eqn. 9.11 is independent of i and depends only on the geometry of the device.

(i) The inductance of a solenoid

Let us consider a section of length l of a long solenoid of cross-sectional area A. It is assumed that the section is near the centre of the solenoid so that edge effects need not be considered. From eqn. 8.22 chap. viii we know that the magnetic field B inside a solenoid carrying a current i is given by

$$B = \mu_0 ni \quad (9.12)$$

where n is the number of turns per unit length. The flux linked with each turn is given by BA. Therefore, the number of flux linkages in the length l is then given by

$$N\Phi_B = (nl)(BA)$$

$$= nl\,\mu_0 ni\,A$$

$$= \mu_0\, n^2 i\, l\, A$$

The self inductance of the coil, as given by eqn. 9.10, is

$$L = \frac{N\Phi_B}{i} = \frac{\mu_0 n^2 ilA}{i} = \mu_0 n^2 lA \quad (9.13)$$

The inductance per unit length of the solenoid is

$$\frac{L}{l} = \frac{\mu_0 n^2 lA}{l} = \mu_0 n^2 A \quad (9.14)$$

If the solenoid is wound on an iron core of relative permeability μ_r, then the inductance (eqn. 9.13) is given by

$$L = \mu_r \mu_0 n^2 lA$$

and the inductance per unit length is then given by

$$\frac{L}{l} = \mu_r \mu_0 n^2 A$$

As can be seen, the expression given by eqn. 9.13 involves only geometrical factors - the cross - sectional area and the number of turns per unit length. The proportionality to n^2 is expected; if the number of turns per unit length is doubled, not only the number N of turns doubled, but the flux Φ_B through each turn is doubled. The number of flux linkages, therefore, increases by a factor 4, as does the inductance.

(i) The inductance of a toroid

Let us now calculate the inductance of a toroid of rectangular cross-section, as shown in Fig. 9.11. The magnetic field in a toroid, as given by eqn. 8.23 chap. viii, is

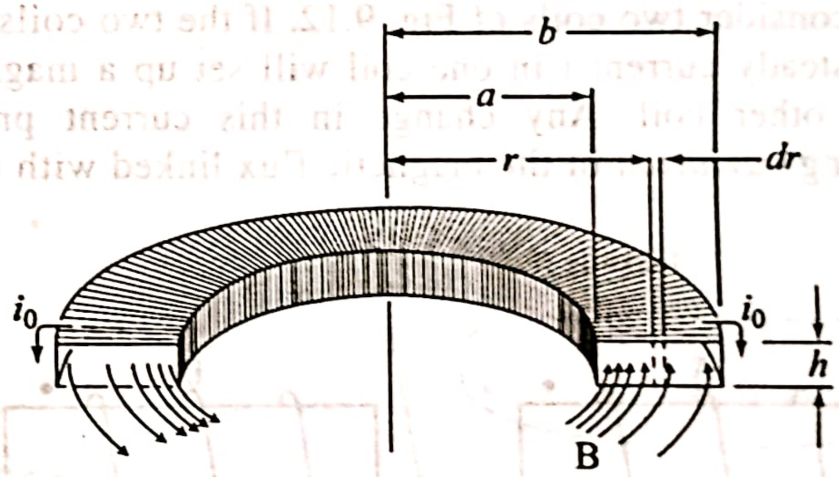


Fig. 9.11

$$B = \frac{\mu_0 i N}{2\pi r}$$

where N is the total number of turns of the toroid. It may be noted that the magnetic field is not constant inside the toroid but varies with the radius r.

The flux ϕ_B through the cross–section of the toroid is

$$\phi_B = \int B.dA = \int_a^b B(hdr) = \int_a^b \frac{\mu_0 iN}{2\pi r} hdr$$

$$= \frac{\mu_0 iNh}{2\pi}\int_a^b \frac{dr}{r}$$

$$= \frac{\mu_0 iNh}{2\pi} \ln\frac{b}{a}$$

where h.dr is the area of the elementary strip between the dashed lines shown in Fig. 9.11. The inductance can then be found from eqn. 9.10.

$$L = \frac{N\phi_B}{i} = \frac{\mu_0 N^2 h}{2\pi} \ln\frac{b}{a} \tag{9.15}$$

Once again, L depends only on geometrical factors.

9.7 Mutual induction

Let us consider two coils of Fig. 9.12. If the two coils are close together, a steady current i in one coil will set up a magnetic flux linking the other coil. Any change in this current produces a corresponding *variation* of the magnetic flux linked with the other.

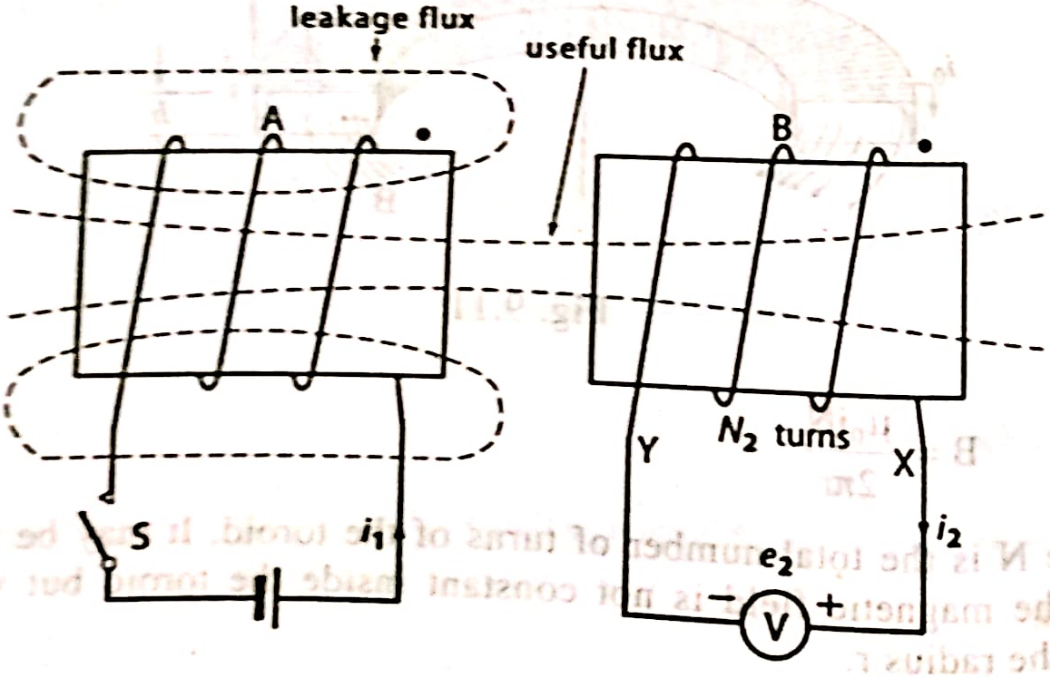


Fig. 9.12

This variation of magnetic flux in the second coil induces a transient emf in it which, by Lenz's law, opposes the growth or decay of the current in the first coil. This induction between the two coils which delays the growth or decay of the current in one due to variation of magnetic flux in the other, is called *mutual induction* to distinguish it from *self-induction*, in which only one coil is involved.

Mutual inductance

Two coils are said to be mutually coupled when a change in the magnetic flux produced by one coil causes an emf to be induced in the other. For this to occur, the flux produced by either coil must *link* or *cut* the windings of the other coil. If the current in one remains constant and the distance between the two coils changes, an emf is produced in the other.

Let us consider the mutually coupled coils of Fig. 9.12. When the switch S is closed, the current in the *primary winding* (coil 1) flows in the direction shown, and while the current is increasing from zero to its final value, the magnetic flux produced by the coil also increases. An emf $\mathcal{E}_2$ is induced in the *secondary winding* (coil 2) *during the time that the flux ϕ_1 in the primary is changing*. The magnitude of the flux linked with the secondary will depend on (i) the strength of the current in the primary, (ii) the number of turns of wire in each coil, (iii) the areas of cross-section of the coils, and (iv) their relative positions. If the factors (ii) (iii) and (iv) are constant, the flux through the secondary will be proportional to (i). According to Faraday's law of electromagnetic induction, the emf induced in the secondary is

$$\mathcal{E}_2 = -N_2\frac{d\phi_2}{dt}$$

where N_2 is the number of turns of the secondary and $d\phi_2$ is the change in flux in the secondary (flux in the secondary is only a part of the flux produced in the primary). Assuming there is no magnetic material inside the secondary, if ϕ_1 is the flux produced in the primary due only to flow of current i_1 in the primary, then

$$\phi_1 \propto i_1$$

But the flux Φ_2 in the secondary is proportional to Φ_1. Hence

$$\Phi_2 \propto i_1$$

Now Φ_2 is the flux linked with each turn of the secondary. Hence the total number of flux linkage with the secondary is $N_2\Phi_2$.

$$\therefore \quad N_2\Phi_2 \propto i_1$$

$$\text{or,} \quad N_2\Phi_2 = Mi_1 \tag{9.16}$$

To be more precise eqn. 9.16 can be rewritten in the form

$$N_2\Phi_{21} = M_{21}i_1 \tag{9.17}$$

where Φ_{21} is the flux through coil 2 associated with the current in coil 1 and M_{21} is the *coefficient of mutual induction* or *mutual inductance* of coil 2 with respect to coil 1. Eqn. 9.17 can be rearranged to obtain an expression for M_{21}. Or,

$$M_{21} = \frac{N_2\Phi_{21}}{i_1} \tag{9.18}$$

Eqn. 9.18 can be compared with the expression for self-inductance $\left(L = \frac{N\Phi}{i}\right)$

If by some external means, i_1 is caused to vary with time, then eqn. 9.18 can be written as

$$M_{21}\frac{di_1}{dt} = N_2\frac{d}{dt}(\Phi_{21}) \tag{9.19}$$

The right side of eqn. 9.19 from Faraday's law is, apart from a difference in sign, is just the emf $\mathcal{E}_2$ appearing in coil 2 due to the changing current in coil 1. Thus

$$\mathcal{E}_2 = -M_{21}\frac{di_1}{dt} \tag{9.20}$$

Again, eqn. 9.20 can be compared with eqn. 9.10 for self-induction $\left(\mathcal{E} = -L.\frac{di}{dt}\right)$.

If the roles of the coils 1 and 2 are now interchanged by setting up a current i_2 in coil 2, by means of a battery, then this will produce a magnetic flux Φ_{12} that links coil 1. If the current i_2 is caused to change, then by the same argument given above, the emf produced in coil 1 is given by

$$\varepsilon_1 = -M_{12}\frac{di_2}{dt} \tag{9.21}$$

where M_{12} is the mutual inductance of coil 1 with respect to coil 2.

Thus it can be seen that the emf induced in either coil is proportional to the rate of change of current in the other coil. The proportionality constants M_{21} and M_{12} seem to be different. It may be asserted, without proof, that they are in fact the same so that no subscripts are needed.

Thus we have

$$M_{12} = M_{12} = M$$

Eqns. 9.20 and 9.21 can therefore, be written as

$$\varepsilon_2 = -M\frac{di_1}{dt} \tag{9.22}$$

and

$$\varepsilon_1 = -M\frac{di_2}{dt} \tag{9.23}$$

The induction is indeed mutual. Rewriting eqns. 9.22 and 9.23 in the form (ignoring the sign)

$$M = \frac{\varepsilon_2}{di_1/dt} = \frac{\varepsilon_1}{di_2/dt}$$

we find that

$$\text{Mutual inductance} = \frac{\text{emf induced in secondary (or primary)}}{\text{rate of change of current in primary (or secondary)}}$$

If $\frac{di_1}{dt} = \frac{di_2}{dt} = 1$, then the mutual inductance can be defined as the emf induced in the secondary (or primary) for a unit rate of change of current in primary (or secondary).

The practical as well as SI unit of mutual inductance is henry. The mutual inductance between two coils is 1 henry, if an emf of 1 volt is induced in the secondary (or primary) when the current in the primary (or secondary) changes at the rate of 1 ampere per second.

Mutual inductance of two co-axial solenoids

The mutual inductance of two solenoids is constant for any pair of coils in a given position provided they do not contain a magnetic material in their core.

Suppose we have two co-axial solenoids of wire closely wound one upon the other. Let the inner one which may be called the primary P, be of length l, cross– sectional area A and of N_1 turns. Let N_2 be the total number of turns of the secondary solenoid S which is wound over the middle of P. If i is the current that flows through P, then the magnetic induction produced by it is

$$B = \mu_0 ni = \mu_0 \frac{N_1}{l} i$$

where $n = \frac{N_1}{l}$ is the number of turns per unit length of the primary.

The flux linked with each turn of the primary is

$$\Phi = B.A = \mu_0 niA$$

The same flux is also linked with each turn of the secondary coil. Hence, the total number of *flux linkage* with the secondary is

$N_2\Phi$ where N_2 is the total number of turns of the secondary.

If the current in the primary is caused to change, then the emf induced in the secondary is (ignoring the sign)

$$\mathcal{E}_2 = \frac{d}{dt}(N_2\Phi) = \frac{d}{dt}(N_2\mu_0 niA)$$

$$= M\frac{di}{dt}$$

or, $M = \mu_0 nN_2A$

$$= \frac{\mu_0 N_1 N_2 A}{l} \quad (9.24)$$

If the solenoid contains an iron core of relative permeability μ_r, then

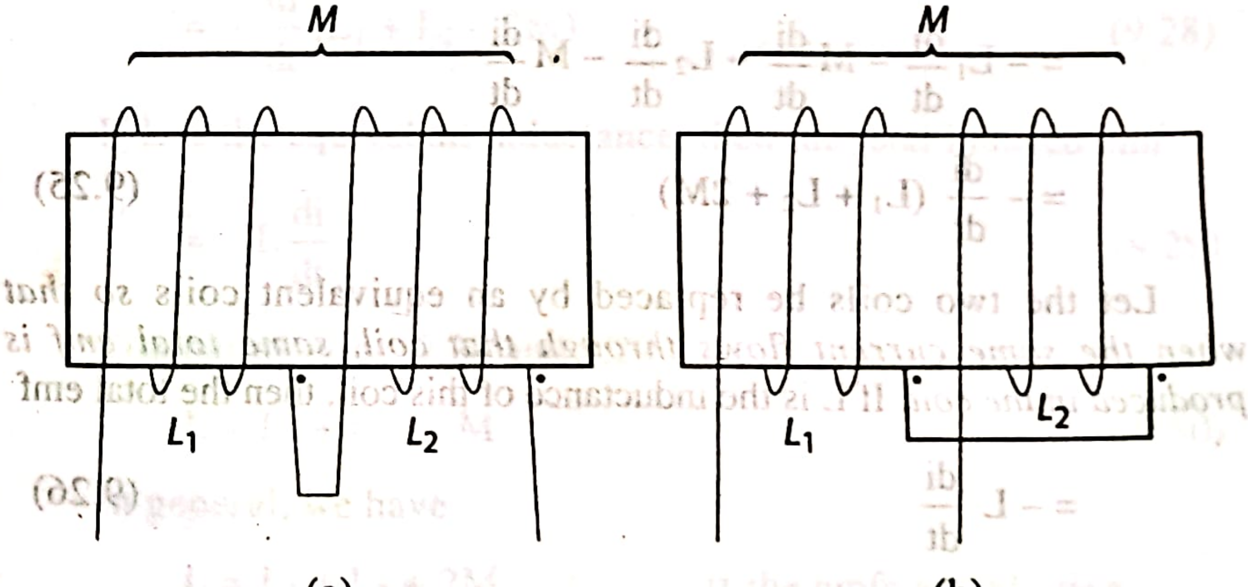
$$M = \frac{\mu_r \mu_0 N_1 N_2 A}{l}$$

9.8 Inductance in series

When two coils wound on the same former are electrically connected together, then the flux that links the coils will modify the value of the induced emf in the coils. In the absence of mutual coupling between the two coils of inductance L_1 and L_2, their inductance when they are connected in series is $L_1 + L_2$; the effect of mutual coupling is to alter the net value of the inductance.

series aiding connection

Let the two coils be so joined in series that the fluxes produced by the coils act in the same direction *i.e.*, fluxes (or emfs) are additive. This sort of connection is called *series aiding* connection [Fig. 9.13(a)]



(a) (b)

Fig. 9.13

Let M = mutual inductance between the coils

L_1 = self inductance of first coil (A)

L_2 = self inductance of second coil (B)

Then self-induced emf in A is

$$e_1 = -L_1\frac{di}{dt}$$

Mutually induced emf in A due to change in current in B

$$e_1' = -M\frac{di}{dt}$$

Self–induced emf in B is

$$e_2 = -L_2\frac{di}{dt}$$

Mutually induced emf in B due to change in current in A

$$e_2' = -M\frac{di}{dt}$$

All the induced emfs have negative sign, because both self and mutually induced emfs are in the same direction and act in opposition to the applied emf.

Thus, the total induced emf in the combination

$$= -L_1\frac{di}{dt} - M\frac{di}{dt} - L_2\frac{di}{dt} - M\frac{di}{dt}$$

$$= -\frac{di}{dt}(L_1 + L_2 + 2M) \qquad (9.25)$$

Let the two coils be replaced by an equivalent coils *so that when the same current flows through that coil, same total emf is produced in the coil.* If L is the inductance of this coil, then the total emf

$$= -L\frac{di}{dt} \qquad (9.26)$$

Comparing eqns. 9.25 and 9.26, we get

$$L = L_1 + L_2 + 2M \qquad (9.27)$$

Series opposing connection

When the coils are so joined that their fluxes are in opposite directions, the connection is called *series opposing* connection [Fig.13(b)]

$$e_1 = -L_1 \frac{di}{dt}$$

$$e_1' = +M \frac{di}{dt} \text{ (the direction is now opposite)}$$

$$e_2 = -L_2 \frac{di}{dt}$$

and $e_2' = +M \dfrac{di}{dt}$

Total induced emf = $e_1 + e_1' + e_2 + e_2'$

$$= -L_1 \frac{di}{dt} + M \frac{di}{dt} - L_2 \frac{di}{dt} + M \frac{di}{dt}$$

$$= -\frac{di}{dt}(L_1 + L_2 - 2M) \quad (9.28)$$

If L is the equivalent inductance, then the total induced emf

$$= -L \frac{di}{dt} \quad (9.29)$$

Comparing eqns. 9.28 and 9.29

$$L = L_1 + L_2 - 2M \quad (9.30)$$

In general, we have

$L = L_1 + L_2 + 2M$ if the emfs are additive.

and $L = L_1 + L_2 - 2M$ if the emfs are subtractive.

or, $L = L_1 + L_2 \pm 2M$

9.9 Coefficient of coupling

If two coils are inductively coupled, in general only a part of the flux arising from one coil is linked with the other coil. If there is very little leakage of flux, the coupling between the two coils is said

to be close or tight, but if the leakage of the flux is large the coupling is said to be *lose*.

The closeness of coupling is expressed by a factor, called the *coefficient of coupling*, which is defined *as the ratio between the mutual inductance and the square root of the product of self inductances* and may be deduced as follows.

Let the coil L_1 produce a flux Φ_1 due to a flow of current i_1 in it. A portion $K\Phi_1$ of it will link with the coil L_2 where K is a constant known as the *coupling coefficient*.

Let the flux linked with L_2 be Φ_1'.

Now $\qquad \Phi_1' = Mi_1$

$$= K\Phi_1$$

Hence $\qquad M = \frac{K\Phi_1}{i_1}$

If we now consider the flow of current i_2 in coil L_2, it can be shown in a similar manner that

$$M = \frac{K\Phi_2}{i_2}$$

Thus $\qquad M \times M = M^2 = \frac{K^2\Phi_1\Phi_2}{i_1 i_2}$

But $\frac{\Phi_1}{i_1} = L_1$ and $\frac{\Phi_2}{i_2} = L_2$ $\qquad [\phi_1 = L_1 i_1$, $\phi_2 = L_2 i_2]$

Hence $\qquad M^2 = K^2 L_1 L_2$; $\qquad$ or, $K^2 = \frac{M^2}{L_1 L_2}$

$$\therefore \quad K = \frac{M}{\sqrt{L_1 L_2}}. \qquad (9.31)$$

9.10 Energy stored in a magnetic Field

If has already been mentioned that the inductance possessed by a conductor or an electric circuit, is analogous to mechanical inertia,

and for this reason it is said to have an electromagnetic inertia, and any arrangement, which increases the flux linked with the circuit, increases its inductance or the inertia effect.

In a mechanical system when a force F is applied to a mass to produce, say, an acceleration or increase in velocity v then according to Newton's second law of motion, $F = m.\frac{dv}{dt}$, where dv is the increase in velocity. To vary a current in a circuit, it requires an emf $e = L\frac{di}{dt}$. Thus L is analogous to m, and $\frac{di}{dt}$ to $\frac{dv}{dt}$. Now any mass resists a change in its velocity, and therefore, some work must be done in giving it a velocity; the work done in giving it kinetic energy $= \frac{1}{2} mv^2$.

Similarly the inductance of a circuit resists any change in the current through it, or establishing the magnetic field, and work is done in increasing the current from o to i_0 against an induced emf. The work done by the current is stored up in the magnetic field. The amount of work W done in this case is given by

$$W = \int_0^{i_0} eidt = \int_0^{i_0} L\frac{di}{dt}idt$$

$$= L\int_0^{i_0} i\frac{di}{dt}dt = L\left[\frac{1}{2}i^2\right]_0^{i_0}$$

If the current flowing in an inductive circuit of inductance L increases at a uniform rate from o to i amperes in a time of 't' seconds, then the *average* circuit current is $\frac{i}{2}$ amperes, and the average value of induced emf is L x rate of change of current or $L\frac{i}{t}$ volts. Hence the *average energy* consumed by the inductive circuit is

$$W = eit = \frac{Li}{t} \times \frac{i}{2} \times t = \frac{1}{2}Li^2 \text{ joules.}$$

9.11 Energy Density and the magnetic Field

We shall now derive an expression for the energy density v_B in a magnetic field. Let us consider a length near the centre of a long solenoid of area of cross-section A. Therefore, the volume associated with this length of the solenoid is Al. The stored energy must lie entirely within this volume because the magnetic field outside such a solenoid is essentially zero. Moreover, since the magnetic field is uniform every where inside the solenoid, the stored energy must be stored uniformly throughout this volume (Al) of the solenoid.

Thus if the energy stored is U_B, then the energy density or energy stored per unit volume is

$$v_B = \frac{U_B}{Al}$$

Now $U_B = \frac{1}{2}Li^2$

$$\therefore \quad v_B = \frac{\frac{1}{2}Li^2}{Al} \tag{9.33}$$

Now L in the above expression can be written as $L = \mu_0 n^2 lA$ (eqn. 9.13). From the expression $B = \mu_0 in$ (for a solenoid, eqn. 8.22), we have

$$i = \frac{B}{\mu_0 n}$$

Substituting these values for L and i in eqn. 9.33

$$U_B = \frac{\frac{1}{2}\mu_0 n^2 lA\left(\frac{B}{\mu_0 n}\right)}{Al}$$

$$= \frac{1}{2}\frac{B^2}{\mu_0} \tag{9.34}$$

Eqn. 9.34 gives the energy density stored at any point (in a vacuum or in a non-magnetic substance) where the magnetic induction is B. Although derived by considering a special case, *i.e.*, a solenoid, the equation is true for all magnetic field configurations.

Eqn. 9.34 can be compared with eqn. 4.17 $\mu_E = \frac{1}{2}\varepsilon_0 E^2$ which gives the energy density (in a vacuum) at any point in an electric field. It may be noted that both u_B and μ_E are proportional to the square of the appropriate field quantity, B or E.

Example 9.9 Calculate the average value of self induced emf in a coil of 0.5H when the current flowing through it is increased from 0.1A to 2.1A in 50 ms.

Soln.

$$\varepsilon = L.\frac{di}{dt} \qquad di = 2.1 - 0.1 = 2A$$

$$dt = 50ms = 50 \times 10^{-3}s$$

$$= \frac{(0.5)(2)}{50\times10^{-3}}$$

$$= 20\ V.$$

Example 9.10 A coil of 500 turns has an inductance of 15 mH. Determine (i) the flux produced in the core when a current of 2A flows in the coil and (ii) the value of the self–induced emf in the coil when the current is changed form +2A to –2A in 10 ms.

Soln.

From eqn. 9.11

$$L = \frac{\text{change in flux linkages}}{\text{change in current}} = \frac{N(\Phi_2 - \Phi_1)}{I_2 - I_1}$$

where Φ_1 and I_1 are the initial values of magnetic flux and current respectively, and Φ_2 and I_2 are the final values. If the initial values of current and flux are zero, then

$$L = N.\frac{\Phi_2}{I_2}$$

therefore

$$\Phi_2 = \frac{LI_2}{N} = \frac{(15\times10^{-3})(2)}{500}\ Wb = 60\ \mu Wb.$$

(ii) Since the current changes from +2A to –2A in a period of 10ms, then

$$di = -2 - (+2) = -4A$$

$$dt = 10 \text{ ms} = 10 \times 10^{-3} s$$

Therefore,

$$\text{The induced emf } = L\frac{di}{dt}$$

$$= \frac{(15\times10^{-3})(-4)}{10\times10^{-3}} = -6V.$$

***Example 9.11** The iron core of a solenoid has a length of 400 mm and a cross–section of 500 mm² and is wound with 1000 turns of wire per metre of length. Compute the inductance of the solenoid, assuming the relative permeability of the iron core (μ_r) to be constant at 500.*

Soln.

Inductance of the solenoid as given by eqn. 9.13

$L = \mu_r\mu_0 n^2 lA$; or, inductance of the solenoid

$$\text{of the iron core} = \frac{L}{l} = \mu_r\mu_0 n^2 A$$

where $\mu_r = 500$, $n = (1000)(0.400) = 400$

$A = 5 \times 10^{-4} m^2$ and $l = 0.40m.$

$$L = \frac{(500)(4\pi\times10^{-7})(400)^2(5\times10^{-4})}{0.400} = 126 \times 10^{-3} H.$$

$$= 126mH.$$

***Example 9.12** Evaluate the inductance of the toroid of rectangular cross-section as shown in Fig. 9.11 Assume N = 10^3, a = 5.0cm, b = 10cm and h = 1.0cm.*

Soln.

The inductance of the given toroid is given by

$$L = \frac{\mu_0 N^2 h}{2\pi} \ln \frac{b}{a}$$

$$= \frac{(4\pi\times10^{-7})(10^3)^2(1.0\times10^{-2})}{2\pi} \ln \frac{10\times10^{-2}}{5\times10^{-2}}$$

$= 1.4 \times 10^{-3}$ Wb/amp

$= 1.4$ mH.

***Example 9.13** The coefficient of mutual inductance between two coils is 8 mH. What emf is induced in the second coil if the current is changing at the rate of 4kA/s in the first coil?*

Soln.

$$\mathcal{E} = M\frac{\Delta i_1}{\Delta t} = (8 \times 10^{-3}\text{H})\ (4000\ \text{A/s})$$

$$= 32\text{V}.$$

***Example 9.14** Two coils are wound on the same iron rod so that the flux generated by one passes through the other also. The primary coil has N_P loops on it and when a current of 2A flows through it, the flux in it is 2.5×10^{-4} Wb. Determine the mutual inductance of the two coils if the secondary coil has N_S loops on it.*

Soln.

Let us assume that the secondary coil is an open circuit. Now

$$\mathcal{E}_S = N_S\frac{\Delta\phi_S}{\Delta t}; \qquad \text{also} \qquad \mathcal{E}_S = M\frac{\Delta i_P}{\Delta t}$$

From the two equations we get

$$M = N_S\frac{\Delta\phi_S}{\Delta i_P}$$

Now $\Delta\phi_S = (2.5 \times 10^{-4} - 0)$ Wb

and $\Delta i_P = (2 - 0)$A

$$\therefore \quad M = N_S\frac{(2.5\times10^{-4} - 0)\text{Wb}}{(2-0)\text{A}} = (1.25 \times 10^{-4}\ N_S)\text{H}.$$

***Example 9.15** The mutual inductance between two coils is 0.2H. If the current in one winding increases from 100 mA to 600 mA in 5 ms, (a) determine the average value of emf induced in the secondary winding during this period of time, and (b) if the secondary is wound with 500 turns, calculate the change of flux linking with the secondary winding.*

$$\mathcal{E}_S = M.\frac{di_1}{dt} \qquad di_1 = (600 - 100) \times 10^{-3}\ A$$

$$dt = 5\ ms = 5 \times 10^{-3} s$$

$$= \frac{(0.2H)(500\times10^{-3}A)}{(5\times10^{-3})s}$$

$$= 20\ V.$$

Also,

$$M = N_S\frac{d\phi_S}{\Delta i_P}; \quad \text{or} \quad d\phi_S = \frac{M.\Delta i_P}{N_S}$$

$$\text{or} \quad \text{flux change} = \frac{(0.2H)(500\times10^{-3}A)}{500}$$

$$= 0.2 \times 10^{-3}\ Wb = 0.2\ mWb.$$

***Example 9.16** A flux of 0.5 mWb is produced by a coil of 900 turns wound on a ring with a current of 3A in it. Calculate (i) the inductance of the coil, (ii) the emf induced in the coil when a current of 5A is switched off, assuming the current to fall to zero in 1 millisecond and (iii) the mutual inductance between the coils, if a second coil of 600 turns is uniformly wound over the first coil.*

Soln.

(i) inductance of the first coil

$$L = \frac{N\phi}{i} = \frac{900\times0.5\times10^{-3}}{3} = 0.15\ H$$

(ii) $\mathcal{E} = L\dfrac{di}{dt} = 0.15 \times \dfrac{(5-0)}{1\times10^{-3}} = 750\ V.$

(iii) $M = \dfrac{N_2\phi_1}{i_1} = \dfrac{(600)(0.5\times10^{-3})}{3} = 0.1\ H.$

The flux through the solenoid due to the primary is

$$\phi = BA = (\mu_r\mu_0 n i_p)\, A = \mu_r\mu_0 i_p\, A\, \frac{2000}{d}$$

This same flux goes through the secondary. We therefore have

$$\mathcal{E}_S = N_S\frac{\Delta\phi}{\Delta t} \quad \text{and} \quad \mathcal{E}_S = M\frac{\Delta i_p}{\Delta t}$$

$$M = N_s\,\frac{\Delta\phi}{\Delta i_p} = N_s\,\frac{(\phi - 0)}{(i_p - 0)}$$

$$= \frac{(50)\,\mu_r\mu_0 i_p A\,(2000/d)}{i_p}$$

$$= \frac{10^5\ \mu_r\mu_0 A}{d}$$

Example 9.18 The combined inductance of two coils connected in series is 0.6H or 0.1H depending on the relative directions of the currents in the coils. If one of the coils when isolated has a self inductance of 0.2H, calculate (a) mutual inductance and (b) coupling coefficient.

Soln.

(i) $L_1 + L_2 + 2M = 0.6\ H$ (i)

$L_1 + L_2 - 2M = 0.1\ H$ (ii)

$4M = 0.5\ H;$ or $M = 0.125\ H.$

(ii) Let $L_1 = 0.2\ H$; Substituting in either (i) or (ii)

$L_2 = 0.15$ H.

$$\therefore \quad k = \frac{M}{\sqrt{L_1 L_2}} = \frac{0.125}{\sqrt{0.2 \times 0.15}} = 0.72$$

Example 9.19 Two coils are wound on a common magnetic circuit and have inductance of 1 H and 0.6H. If the coupling coefficient is 0.5, determine (i) the mutual inductance between the coils and (ii) the inductance of the circuit if they are connected in (a) series–aiding, (b) series opposing. If the current in the circuit in cases (a) and (b) above changes at the rate of 100 A/s, determine the induced emf in each case.

Soln.

(i) $k = \dfrac{M}{\sqrt{L_1 L_2}}$ $\qquad \therefore \; M = k\sqrt{L_1 L_2} = 0.5 \times \sqrt{0.64 \times 1}$

$= 0.4$ H.

(ii) (a) For series-aiding coils

$$L = L_1 + L_2 + 2M$$

$$= 1 + 0.64 + (2 \times 0.4) = 2.44 \text{ H.}$$

(b) For series-opposing coils

$$L = L_1 + L_2 - 2M$$

$$= 1 + 0.64 - (2 \times 0.4)$$

$$= 0.84 \text{ H.}$$

when the current changes at the rate of 100 A/s the induced emf for

(a) series-aiding coils

$$\mathcal{E} = L.\frac{di}{dt} = 2.44 \times 100 = 244 \text{ V.}$$

(b) series-opposing coils

$$\mathcal{E} = L.\frac{di}{dt} = 0.84 \times 100 = 84 \text{ V.}$$

Example 9.20 A coil has an inductance of 5.0 henrys and a resistance of 20 ohms. If a 100 volt emf is applied, what energy is stored in the magnetic field after the current has built up to its maximum value $\mathcal{E}/R$.

Soln.

The maximum current is given by

$$i = \frac{\mathcal{E}}{R} = \frac{100 \text{ volts}}{20 \text{ ohms}} = 5.0 \text{ amp}$$

The stored energy is given by

$$U_B = \frac{1}{2}Li^2 = \left(\frac{1}{2}\right)(5.0)\,(5.0)^2 = 63 \text{ joules.}$$

Example 9.21 An electromagnet has stored 648J of magnetic energy when a current of 9A exists in its coils. What average emf is induced if the current is reduced to zero in 0.45s?

Soln.

$$\frac{1}{2}Li^2 = 648 \text{ J}$$

$$\frac{1}{2}L\,(9A)^2 = 648 \text{ J}; \text{ or } \quad L = 16H.$$

$$\text{emf induced} = L\frac{\Delta i}{\Delta t} \qquad \Delta i = (9-0)A = 9A$$

$$\Delta t = 0.45s$$

$$\therefore \quad \mathcal{E} = (16\text{ H})\frac{9A}{0.45s} = 320 \text{ V.}$$

Example 9.22 (a) In a certain coil a flux linkage of 3.5 weber-turns is established by a current of 10A. Determine the inductance of the coil. (b) In (a), i is made to decrease at the rate of 100A/s. What is the voltage of self-induction and what is its direction? (c) For a current of 15A, how much energy is stored in the coil?

Soln.

(a) We have $N\Phi = 3.5$ Wb

where N is the total number of turns and Φ is the flux linked with each turn. Hence from $L = N\Phi/i$, we have

$$L = \frac{3.5\ \text{Wb}}{10\ \text{A}} = 0.35\ \text{H}.$$

(b) The emf induced is

$$\mathcal{E} = -L\frac{di}{dt} = -(0.35)(-100) = +35\ \text{V}.$$

Therefore the voltage induced is in the direction of the current.

(c) Stored energy,

$$U_B = \frac{1}{2}Li^2 = \frac{1}{2}(0.35)(15)^2$$

$$= 39.375\ \text{J}.$$

***Example 9.23** A toroid of 0.5m circumference and 480–mm^2 cross-sectional area has 2500 turns bearing a current of 0.6A. It is wound on an iron-ring with relative permeability of 350. Find the magnetizing field H, the magnetic induction B, the flux Φ, the coefficient of self inductance L, and the energy stored in the magnetic field.*

Soln.

(a) The intensity of magnetizing field, H, is given by

$$H = nI$$

where n is the number of turns per unit length and I is the current.

$$\therefore \quad H = nI = \left(\frac{2500}{0.5}\right)(0.6) = 3000\ \text{A/m}.$$

(b) The magnetic induction B is given by

$$B = \mu_r\mu_0 ni$$

$$= (350)(4\pi \times 10^{-7})\left(\frac{2500}{0.5}\right)(0.6)$$

$$= 1.32\ \text{T}.$$

(c) The flux Φ is given by

$$\Phi = B.A. = (1.32\ Wb/m^2)\ (480 \times 10^{-6}\ m^2)$$

$$= 633 \times 10^{-6}\ Wb = 633\mu Wb.$$

(d) The coefficient of self inductance is given by the relation

$$L = \frac{N\Phi}{i} = \frac{(2500)(633\times10^{-6})}{0.6}$$

$$= 2.64\ H.$$

(e) Energy stored,

$$U_B = \frac{1}{2}Li^2 = \left(\frac{1}{2}\right)(2.64)\ (0.6)^2$$

$$= 0.475\ J.$$

9.12 Growth and decay of current in a circuit containing inductance and resistance (LR circuit)

If has already been mentioned that self–inductance in an electric circuit delays the growth and decay of current at make and break of a circuit, because it opposes changes of current . Let us find the rate at which it happens.

(a) Growth of current: Let us consider a circuit having an inductance L and a resistance R placed in series with a cell of emf $\mathcal{E}$ (Fig. 9.14). When the current is started in the circuit by throwing the switch S to a in Fig. 9.14 the current slowly increases from zero to a maximum value i_0. During the time the current is growing a back emf is developed in the inductance. Let i be the value of the current

An *LR* circuit.

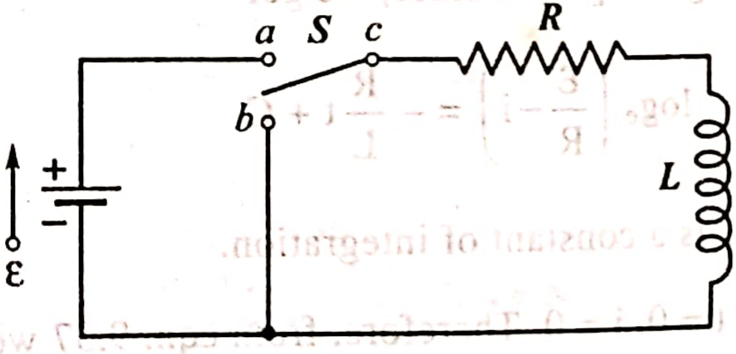


Fig. 9.14

at any instant (instantaneous value) during this variable state and $\frac{di}{dt}$ be the rate of growth of current in the circuit. The fall of potential across the resistance is Ri. The back emf developed in the inductance $= -L\frac{di}{dt}$. Since the back emf opposes the emf of the battery, the equation of emf of the circuit during the growth of the current is given by

$$\mathcal{E} - L\frac{di}{dt} = Ri$$

$$\text{or,} \quad L\frac{di}{dt} + Ri = \mathcal{E} \tag{9.35}$$

Eqn. 9.35 may be written as

$$L\frac{di}{dt} = \mathcal{E} - Ri$$

$$\text{or,} \quad \frac{L}{R}\frac{di}{dt} = \frac{\mathcal{E}}{R} - i$$

$$\text{or,} \quad \frac{L}{R}di = \left(\frac{\mathcal{E}}{R} - i\right)dt \tag{9.36}$$

Eqn. 9.36 may be rearranged as

$$-\frac{di}{\left(\frac{\mathcal{E}}{R} - i\right)} = -\frac{R}{L}dt$$

Integrating both sides, we get

$$\log_e\left(\frac{\mathcal{E}}{R} - i\right) = -\frac{R}{L}t + C \tag{9.37}$$

where C is a constant of integration.

At t = 0, i = 0. Therefore, from eqn. 9.37 we get

$$C = \log_e \frac{\varepsilon}{R}$$

Eqn. 9.37 then becomes

$$\log_e \left(\frac{\varepsilon}{R} - i\right) = -\frac{R}{L}t + \log_e \frac{\varepsilon}{R}$$

$$\text{or,} \quad \log_e \left(\frac{\varepsilon/R - i}{\varepsilon/R}\right) = -\frac{R}{L}t$$

$$\text{or,} \quad \frac{\varepsilon/R - i}{\varepsilon/R} = e^{-\frac{R}{L}t}$$

$$\text{or,} \quad \frac{\varepsilon}{R} - i = \frac{\varepsilon}{R} e^{-\frac{R}{L}t}$$

$$\text{or,} \quad i = \frac{\varepsilon}{R}\left(1 - e^{-\frac{R}{L}t}\right) \qquad (9.38)$$

Eqn. 9.38 gives the value of the current in time t from the instant of closing the circuit. Now $\frac{\varepsilon}{R}$ is the final steady or maximum value i_0 of the current when $L\frac{di}{dt} = 0$.

$$\therefore \quad i = i_0\left(1 - e^{-\frac{R}{L}t}\right) \qquad (9.39)$$

The rate of growth of current is given by

$$\frac{di}{dt} = \frac{R}{L} i_0 \, e^{-\frac{R}{L}t} \qquad (9.40)$$

Eqn. 9.39 states that at any time t the current i is the maximum steady current i_0 multiplied by the diminishing factor $\left(1 - e^{-\frac{R}{L}t}\right)$

which runs from o to 1. If t is indefinitely great $e^{-\frac{R}{L}t}$ tends to zero. If t is indefinitely small, $e^{-\frac{R}{L}t}$ tends to 1.

From eqn. 9.39 it is obvious that the rise of current in an LR circuit is *exponential* as shown in Fig. 9.15. It will, therefore, need an infinite time for the current to attain its final or maximum value. It may further be noted that, the smaller the value of L, the shorter will be the time for the current to attain its maximum value.

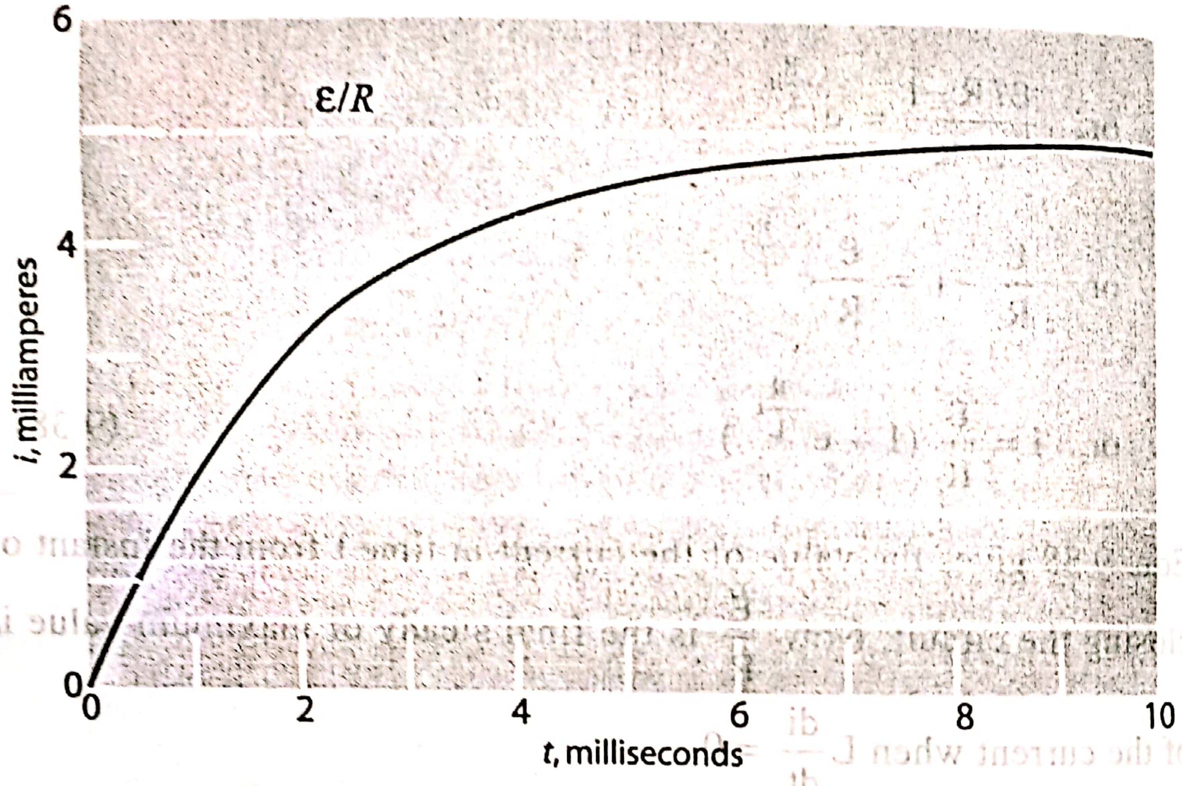


Fig. 9.15

Time constant : Eqn. 9.39 can be rearranged as

$i = i_0 (1 - e^{-\frac{1}{L/R}t})$. The SI unit of inductance is Henry and that of the resistance is ohm. Thus

$$\frac{L}{R} = \frac{1 \text{ henry}}{1 \text{ ohm}} = \frac{1 \text{ volt second/ampere}}{1 \text{ volt/ampere}} = 1 \text{ second.}$$

Therefore the ratio of L/R has dimension of time and is called the *time constant* of the circuit. If in eqn. 9.39, t is numerically equal to L/R, then the value of the rising current is

$$i = i_0 (1 - e^{-\frac{1}{L/R} \cdot t}) = i_0 (1 - e^{-1})$$

$$= i_0 (1 - \frac{1}{2.718}) = 0.632\, i_0 \sim \frac{2}{3}\, i_0$$

Thus L/R or the time constant of a circuit containing inductance and resistance is the time taken for the current to rise from zero to 0.632 or 63% of its final or steady value.

In a *purely inductive circuit* (R = 0), the value of the time constant is infinity. In such a circuit the current takes an infinitely long time to attain even $\frac{2}{3}$rd of its maximum value.

In a *purely resistive circuit* (L = 0), the value of the time constant is zero. This means the current grows very quickly in a purely resistive circuit.

Decay of current in an L-R circuit

If the switch S in Fig. 9.14, having been left in position *a* enough for the equilibrium current $\mathcal{E}/R$ to be established, is thrown to *b*, the effect is to remove the battery from the circuit *i.e.*, $\mathcal{E} = 0$. An induced emf will again be produced. Due to this back emf the current decreases slowly to zero. The equation for the emf may therefore be written for the decay of current as

$$-L\frac{di}{dt} = R\,i \tag{9.41}$$

This equation may be written as

$$\frac{di}{i} = -\frac{R}{L}dt$$

Integrating, we get

$$\log_e i = -\frac{R}{L}t + C \tag{9.42}$$

where C = constant of integration

At $t = 0,\ i = i_0 = \frac{\mathcal{E}}{R}$ (i_0 is the maximum current at the start and $\mathcal{E}$ is the applied emf just before the circuit was switched off)

Therefore $C = \log_e i_0$

Hence eqn. 9.42 may be written as

$$\log_e i = -\frac{R}{L}t + \log_e i_0$$

$$\text{or,} \quad \log_e \frac{i}{i_0} = -\frac{R}{L}t$$

$$\text{or,} \quad \frac{i}{i_0} = e^{-\frac{1}{L/R}.t}$$

$$\text{or,} \quad i = i_0 e^{-\frac{1}{L/R}.t} \tag{9.43}$$

$$= \frac{\varepsilon}{R} e^{-\frac{1}{L/R}.t} \tag{9.44}$$

Eqn. 9.44 gives the value of current at any instant. As can be seen, the decay of current is also exponential as shown in Fig. 9.16. The ratio $\frac{L}{R}$ is again called the time constant of the circuit.

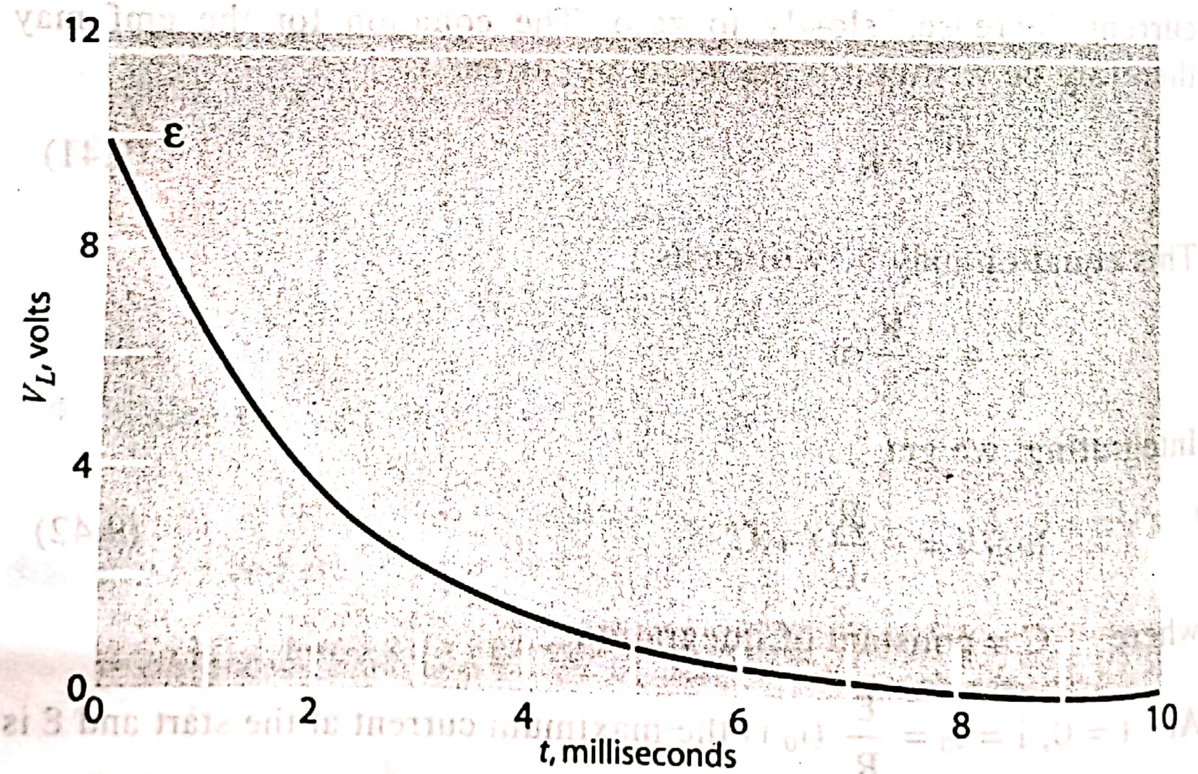


Fig. 9.16

If $t = \frac{L}{R}$, $i = i_0e^{-1} = \frac{i_0}{2.718} = 0.37\ i_0$.

Thus the time constant signifies the time in which the current drops to $\frac{1}{e}$ or approximately $\frac{1}{3}$rd of its initial maximum steady value.

Thus a current reaches 0.632 of its maximum value in 1 second when *starting up* and falls to 0.37 of its maximum value in 1 second when *collapsing*.

If may be noted that when the applied emf is switched off, the collapsing magnetic field in the coil becomes the source of energy for the decaying current.

***Example 9.24** A coil of resistance 15 Ω and inductance 0.06H is connected to a steady 120-V power source. At what rate will the current in the coil rise (a) at the instant the coil is connected to the power source, and (b) at the instant the current reaches 80 percent of its maximum vale?*

Soln.

The effective driving voltage in the circuit is the 120–V power supply minus the induced back emf, $L\left(\frac{\Delta i}{\Delta t}\right)$. This equals the potential drop in the resistance of the coil. Or,

$$120\ V - L\frac{\Delta i}{\Delta t} = Ri$$

(a) At the first instant, i is essentially zero. Then

$$\frac{\Delta i}{\Delta t} = \frac{120\ V}{L} = \frac{120\ V}{0.6\ H} = 200\ A/s.$$

(b) When the current reaches its maximum value E/R, it finally stops changing *i.e.*, $\Delta i/\Delta t = 0$. We are interested in the case when

$$i = 0.8\frac{120\ V}{R}$$

Then substitution in the loop equation gives

$$120\ V - L\frac{\Delta i}{\Delta t} = 0.8\ (i)\ R$$

$$= 0.8\ (\frac{120\ V}{R})\ R$$

or, $L\frac{\Delta i}{\Delta t} = (0.2)\ (120\ V)$

or, $\frac{\Delta i}{\Delta t} = \frac{(0.2)(120\ V)}{0.6\ H} = 40\ A/s.$

***Example 9.25** A coil has an inductance of 53 mH and resistance of 0.35Ω (a) If a 12–V emf is applied, how much energy is stored in the magnetic field after the current has built up to its maximum value? (b) In terms of τ_L, how long does it take for the stored energy to reach half of its maximum value?*

Soln.

(a) Maximum current is given by

$$i_0 = \frac{\varepsilon}{R} = \frac{12\ V}{0.35\ \Omega} = 34.3\ A.$$

Energy stored in the magnetic field,

$$U_B = \frac{1}{2}Li_0^2 = \left(\frac{1}{2}\right)(53 \times 10^{-3}\ H)\ (34.3A)^2 = 31\ J$$

(b) $i = i_0\left(1 - e^{-\frac{R}{L}t}\right) = i_0\left(1 - e^{-\frac{t}{\tau_L}}\right) \qquad \tau_L = \frac{L}{R}$

where i is the current at the instant the stored energy has half its maximum value.

Now

$$\frac{1}{2}Li^2 = \frac{1}{2}\left(\frac{1}{2}Li_0^2\right)$$

or, $i = \frac{i_0}{\sqrt{2}}$

But $i_0 = \frac{\varepsilon}{R}$

$\therefore \quad i = \frac{\varepsilon}{\sqrt{2}R}$

$$\therefore \quad \frac{\varepsilon}{\sqrt{2}R} = \frac{\varepsilon}{R}\left(1 - e^{-\frac{t}{\tau_L}}\right)$$

$$\text{or,} \quad e^{-\frac{t}{\tau_L}} = 1 - \frac{1}{\sqrt{2}} = 0.293$$

$$\text{or,} \quad -\frac{t}{\tau_L} = \ln 0.293 = -1.23$$

$\therefore \quad t = 1.23\,\tau_L$

Example 9.26 A 3.56-H inductor is placed in series with a 12.8Ω resistor. An emf of 3.24V is suddenly applied to the combination. At 0.278s (which is one inductive time constant) after the contact is made find (a) the rate P at which energy is being delivered by the battery (b) the rate P_R at which internal energy appears in the resistor and (c) the rate P_B at which energy is stored in the magnetic field.

Soln.

(a) The current is given by

$$i = i_0\left(1 - e^{-\frac{R}{L}t}\right) = i_0\left(1 - e^{-\frac{t}{\tau_L}}\right) \qquad \tau_L = \frac{L}{R}$$

$$= \frac{\varepsilon}{R}\left(1 - e^{-\frac{R}{L}t}\right) = \frac{3.24\text{ V}}{12.8\,\Omega}(1 - e^{-1}) \qquad t = \tau_L$$

$$= 0.1600 \text{ A}.$$

The rate at which the battery delivers energy is

$$P = \mathcal{E}i = (3.24 \text{ V})(0.16 \text{ A}) = 0.5184 \text{ W}.$$

(b) The rate P_R at which energy is dissipated in the resistor is

$$P_R = i^2R = (0.16 \text{ A})^2 (12.8 \ \Omega) = 0.3277 \text{ W}.$$

(c) The rate at which energy is stored in the magnetic field is given by

$$\frac{d}{dt}(U_B) = \frac{d}{dt}\left(\frac{1}{2}Li^2\right) = Li\frac{di}{dt}$$

Now $$\frac{di}{dt} = \frac{d}{dt}\left[\frac{\mathcal{E}}{R}\left(1-e^{-\frac{t}{\tau_L}}\right)\right]$$

$$= \frac{\mathcal{E}}{R}\cdot\frac{1}{\tau_L}e^{-\frac{t}{\tau_L}} = \frac{\mathcal{E}}{R}\cdot\frac{R}{L}e^{-\frac{t}{\tau_L}} = \frac{\mathcal{E}}{L}e^{-\frac{t}{\tau_L}}$$

At $t = \tau_L$,

$$\frac{di}{dt} = \frac{\mathcal{E}}{L}e^{-1} = \frac{3.24 \text{ V}}{3.56 \text{ H}}e^{-1} = 0.3348 \text{ A/s}.$$

∴ the desired rate

$$P_B = \frac{d}{dt}(U_B) = Li\frac{di}{dt}$$

$$= (3.56 \text{ H})(0.1600 \text{ A})(0.3348 \text{ A/s})$$

$$= 0.1907 \text{ W}.$$

It may be noted that as required by principle of conservation of energy.

$$P = P_R + P_B = 0.3277 \text{ W} + 0.1907 \text{ W}$$

$$= 0.5184 \text{ W}.$$

EXERCISES

1. State and explain laws of electromagnetic induction.

2. State and explain Lenz's law.

3. What is motional electromotive force?

   A close conducting loop is caused to move with a uniform velocity with respect to a uniform magnetic field. Obtain expressions (i) for the emf induced in the loop, (ii) the rate of work done by the agent that causes the loop to move and (iii) the rate at which internal energy is produced by the current.

4. Show that the rate at which heat appears across a closed rectangular loop of width l when it is pulled with respect to a uniform magnetic field of induction B with a constant velocity v is given by $\frac{B^2l^2v^2}{R}$ where R is the resistance of the loop.

5. Explain the terms self inductance and mutual inductance.

6. Explain self-inductance. Show that the dimensions of the two expressions for self inductance $L = \frac{N\Phi_B}{i}$ and $L = \frac{\varepsilon}{di/dt}$ are the same.

7. Show that the energy required to build up a current i in a circuit having self–inductance L is $\frac{1}{2}Li^2$.

8. State and explain Faraday's law of electromagnetic induction and Lenz's law for the direction of induced emf.

9. Define self inductance, mutual inductance and co-efficient of coupling Show that the coefficient of coupling between two coils is given by $k = M\sqrt{L_1L_2}$ where the symbols have their usual meanings.

10. Two coils of inductance L_1 and L_2 are connected in series. If M is the mutual inductance between the coils, show that their equivalent inductance is given by $L = L_1 + L_2 \pm 2M$ depending on whether the connection is series aiding or opposing.

11. Define Self –inductance. Obtain an expression for the inductance of a solenoid per unit length near the centre of the solenoid in terms of the number of turns per unit length of the solenoid and area of cross-section of the solenoid. Will the inductance per unit length of the

solenoid near its ends be (i) same as (ii) greater then and (iii) less then the inductance per unit length at the centre of the solenoid?

12. Define mutual-inductance.

    Obtain an expression for the mutual inductance of two co-axial solenoids closely wound one upon the other.

13. A circuit contains an inductance L and a resistance R placed in series with a cell of emf $\mathcal{E}$. Obtain expression for the growth and decay of current in the circuit. What is time constant of the circuit?

14. Obtain an expression for the energy density of the magnetic field produced within a solenoid when a current passes through it.

15. Show that when a closed conducting loop is moved across a magnetic field, an induced current is produced in the loop which produces internal energy at the same rate at which mechanical work is done on the loop.

16. The magnetic flux through a spark coil of 1000 turns changes form 0.5 Wb to zero in 0.01 sec. Determine the emf induced in the coil. **[50 KV]**

17. A solenoid 600 mm long has 5000 turns on it and is wound on an iron rod of 7.5mm radius. Find the flux through the solenoid when the current in it is 3A.If this flux is reduced to a value of 1 mWb in a time of 50ms, find the induced emf in the solenoid. **[1.66 mWb, 66 V]**

18. A flat coil with radius 8mm has 50 loops of wire on it. It is placed in a magnetic field, B = 0.30T, so that the maximum flux goes through it. Later, it is rotated in 0.02s to a position such that no flux goes through it. Find the average emf induced between the terminals of the coil. **[0.15 V]**

19. A coil of 275 turns with an area of 0.024m^2 is placed with its plane perpendicular to the earth's field and is rotated in 0.25s through a quarter turn, so that its plane is parallel to the earth's field. What is the average emf induced if the earth's field has an intensity of 80μT? What was the original flux through each turn? **[0.0211 V, 1.92 μWb]**

20. The secondary of an induction coil has 12,000 turns. If the flux linking the coil changes from 740 to 40μWb in 180 μs, how great is the induced emf? **[46.7 KV]**

21. An emf of 3.5 V is obtained by moving a wire 1.1m long at a rate of 7m/s perpendicular to the wire and to a uniform magnetic field. What is the intensity of the field? [0.455 T]

22. An axle of truck is 2.4m long. If the truck is moving due north at 30m/s at a place where the vertical component of earth's magnetic field is 90μT, what is the potential difference between the two ends of the axle? **[6.48 mV]**

23. A section of a solenoid of length $l=12$cm and having a circular cross–section of diameter $d=1.6$cm carries a steady current of $i = 3.80$A. The section contains 75 turns along its length. (a) What is the inductance of the solenoid when the core is empty? (b) The current is reduced at a constant rate to 3.20 A in a time of 15 s. What is the resulting emf developed by the solenoid, and in what direction does it act? (c) The core of the solenoid is filled with iron, If the current is held constant at 3.20 A, the magnetization of the iron is saturated such that $B=1.4$T. What is the resulting inductance? **[12 μH, 0.48 μV, 6.7 mH]**

24. The current in a circuit changes from 24 A to zero in 3ms. If the average induced emf is 260 V, what is the coefficient of self-inductance of the circuit? How much energy was initially stored in the magnetic field of the inductor? **[32.5 mH, 9.36 J]**

25. What back emf is induced in a coil of self-inductance 0.008H when the current in the coil is changing at the rate of 110A/s? What energy is stored in the inductor when the current is 6A? **[0.88 V, 0.144 J]**

26. When the current in a certain coil is changing at a rate of 3A/s it is found that an emf of 7mV is induced in a nearby coil. What is the mutual inductance of the combination? **[2.33 mH]**

27. When the current in the primary of a small transformer is changing at a rate of 600 A/s, the induced emf in the secondary is 8V. What is the coefficient of mutual inductance? **[13.3 mH]**

28. Two neighbouring coils, A and B, have 300 and 600 turns, respectively. A current of 1.5A in A causes 1.2×10^{-4} Wb to pass through A and 0.9×10^{-4} Wb through B. Determine (a) self-inductance of A, (b) the mutual inductance of A and B, and (c) the average induced emf in B when the current in A is interrupted in 0.2S. **[24 mH, 36 mH, 0.27 V]**

29. A 20 mH coil is connected in series with a 2 KΩ resistor, a switch, and a 12-V battery. What is the time constant of this circuit? How long after the switch is closed will it take for the current to reach 99 percent of its final value? **[10 μs, 46 μs]**

30. An inductor L, a resistance R and an emf = V_0 are connected in series with a switch. The switch is pushed closed at t = 0. Find the following: (a) I at t = 0; (b) induced emf at t = 0 and (c) at $t = \frac{1}{2}(L/R)$ (d) I at $\frac{1}{2}(L/R)$.

[0, V_0, 0.607 V_0, (0.393 V_0)/R]

CHAPTER - X

MAGNETISM

We have seen in the previous chapters that an electric current is associated with magnetic field and a current loop is equivalent to a magnetic shell of magnetic moment IA. We also know that a piece of iron bar can easily be magnetized either by rubbing the iron bar with another permanent magnet or by passing electric current through an isolated coil wound around the iron bar. Whatever may be the source of magnetic field (a current loop, a solenoid or a bar magnet), the magnetic field strength at any point is always due to magnetic dipoles of the sources.

There is an important difference between electric and magnetic dipoles: an electric dipole can be separated into its constituent single charges (or poles), but a magnetic dipole cannot. All attempts to isolate a single north pole or a south pole fail. If a magnet is broken we end up with smaller magnets each with its north and south poles. This breakup process of the magnet can be pushed as far as its constituent atoms and electrons but we still fail to find anything that we can call an isolated magnetic pole or a *magnetic monopole* as we have come to call it. Thus the simplest magnetic structure that can exist in nature is the magnetic dipole. There are no magnetic monopoles; that is, there are no magnetic structures analogous to isolated electric charges. The smallest unit – the unit which retains the properties of a magnet – is called a *domain*.

The magnetic property of a material is, therefore, associated with the intrinsic magnetic moment of the constituent dipoles. If these dipoles are arranged in a suitable way either permanently or temporarily by any means the material would exhibit magnetic properties.

10.1 Magnetization

The phenomenon of making a piece of unmagnetized substance into a magnet by the influence of another magnet (or magnetic field) with or without actual contact is referred to as *magnetic induction.* When such a magnetizable substance is placed in a magnetic field, then we will have two sets of lines of force within this magnetizable

substance – (1) those which are due to the magnetizing field and which would exist if the magnetic substance were replaced by a non-magnetic substance; and (2) those due to the magnetization of the substance itself when placed in the magnetic field. The latter are called *lines of magnetization.*

Let a cylindrical bar of a magnetic substance of area of cross-section A and length l, is placed inside a solenoid with its length parallel to the axis of the solenoid. A current i is passed through the solenoid. A magnetic field is produced threading through the magnetic substance. The magnetic substance is composed of atoms having magnetic dipole moments μ_i. These dipoles in general point in various directions in space. The net dipole moment μ of a volume V of the material can be computed by taking the *vector* sum of all the dipoles in that volume: $\mu = \sum \mu_i$. The field inside the solenoid (magnetizing field) *magnetizes* the material and aligns the dipoles in a particular direction. The aligned dipoles produce a magnetic field of their own. The extent of magnetization of the material is expressed by the total magnetic moment per unit volume of the material. This ratio is called the *intensity of magnetization* or simply *magnetization*, M. Hence

$$M = \frac{\mu}{V} = \frac{ml}{Al} = \frac{m}{A} \qquad (10.1)$$

where m = pole strength and l = length.

M is also expressed as number of lines of force per unit area.

Besides the magnetic lines of force inside the material there are magnetic lines of force produced by the solenoid coil, the number of which per unit area = H. H is called the *intensity of the magnetizing field* and is due to the current passing through the coil. Thus we have two sets of lines of fore both in the same direction through the material *i.e.*, (1) due to polarization of the dipoles of the material resulting in I_m and (2) due to the solenoid producing the magnetizing field H, the induction due to its magnetization in empty space = $\mu_0 M$. The net magnetic field B or the total number of lines of force per unit area (or lines of induction) at any point in space is the sum of the applied field B_0 and that produced by the dipoles, which we call B_M, so that the net field can be written as

$$B = B_0 + B_M$$

$$= \mu_0 H + \mu_0 M$$

$$= \mu_0 (H + M) \qquad (10.2)$$

where μ_0 = permeability of free space. If the magnetic material is absent then $B = B_0 = \mu_0 H$ because there are no dipoles to be aligned.

The net field is then given by

$$B = \mu_0 H + \mu_0 M = B_0 + \mu_0 M \qquad (10.3)$$

In weak fields, M increases linearly with the applied field B_0 and so B must be proportional to B_0. In this case, can write

$$B = k_m B_0 \qquad (10.4)$$

where k_m is the *absolute permeability constant* of the material, which is defined relative to a vacuum for which $k_m = 1$. Combining equations 10.3 and 10.4, we can write

$$k_m B_0 = B_0 + \mu_0 M$$

$$\text{or, } \mu_0 M = (k_m - 1) B_0 \qquad (10.5)$$

Eqn. 10.5 gives the magnetization induced by the applied field.

The total number of lines of force crossing any area at right angles to the surface is called *total induction* or *magnetic flux*, and is generally represented by Φ; B is the flux per unit area. If A is the area of the surface and B is the induction, Φ = BA. The unit of magnetic flux density is 1 Weber and the unit of magnetic flux density (or magnetic induction) is 1 Tesla or 1 Weber/meter^2. The previous units for magnetic flux and magnetic flux density were a maxwell and a gauss respectively where 1 Weber = 10^8 maxwells and 1 Tesla =10^4 gauss.

Magnetic intensity and magnetic induction should not be confused; the former has to do with the property of the field expressed in terms of force while the latter with the condition of the medium.

Just as H describes the magnetic (or magnetizing) field at a point, B describes the field at the point provided in this case the presence of the medium is included in the expression for B.

Since the intensity of the magnetic field at a point is measured by the number of lines of force crossing unit area surrounding that point or by the magnetic flux round that point, we find that

H is the flux density before and B is the flux density after the introduction of the material. Each measures the force on unit pole, the former in the absence of the magnetic material, while the latter inside the magnetic material

H, M and B are vector quantities.

In the case of a permanent magnet, since there is no inducing field, $B = \mu_0 M$.

The following point in connection with B and H should be noted: *H is the cause and B is the effect*

If a ferromagnetic core is inserted inside the former of an air-cored coil the flux density is intensified, the factor by which it is increased is given by the ***relative permeability***, μ_r of the material where

$$\mu_r = \frac{\text{flux density with the ferromagnetic core}}{\text{flux density without the ferromagnetic core}}$$

$$= \frac{k_m H}{\mu_0 H} = \frac{k_m}{\mu_0}$$

or, $k_m = \mu_0 \mu_r$

Thus,

Absolute permeability of a material

= relative permeability x permeability of free space.

Or,

$B = \mu_0 \mu_r H$ ($\mu_0 \mu_r M$)

10.2 Magnetic susceptibility and permeability

The intensity of magnetization M, acquired by a magnetic substance depends on the strength of the magnetizing field, the nature and condition of the substance. If H is the strength of the inducing field,

M is the resulting intensity of magnetization, the ratio $\frac{M}{H}$ measures the *magnetic susceptibility* (k) of the material, *i.e.*,

$$k = \frac{M}{H}, \quad \text{or} \quad M = kH$$

Thus the susceptibility of a material is measured by the intensity of magnetization induced in it by unit magnetizing field. The magnetic susceptibility is a *measure of the ease* with which substance is magnetized and varies with the magnitude of the magnetizing field. The ratio of the magnetic induction B produced in a material to the magnetizing field H (in the absence of any magnetic material) is called **absolute permeability** (or magnetic space constant) μ of the material, *i.e.*,

$$\mu = \frac{B}{H}, \quad \text{or} \quad B = \mu H = \mu_0\mu_r H.$$

In vacuum $B = \mu_0 H$ in teslas or Weber/m^2.

***Example 10.1** The magnetic field in the interior of a certain solenoid has the value 6.5×10^{-4}T when the solenoid is empty. When it is filled with iron, the field becomes 1.4 T (a) Find the relative permeability under these conditions. (b) Find the average magnetic moment of an iron atom under these conditions.*

Soln.

(a) From eqn. $\mathbf{B} = \mu_r \mathbf{B_0}$, we have (taking magnitudes only)

$$\mu_r = \frac{B}{B_0} = \frac{1.4T}{6.5\times10^{-4}T} = 2300$$

(b) Using eqn. $\mathbf{B} = \mathbf{B_0} + \mu_0\mathbf{M}$, we get (again taking magnitudes only)

$$M = \frac{B - B_0}{\mu_0} = \frac{1.4T - 6.5\times10^{-4}T}{4\pi\times10^{-7}T.m/A}$$

$$= 1.11 \times 10^6 \text{ A/m (or A.m}^2\text{/m}^3\text{)}$$

This represents the magnetic moment per unit volume of the iron. To find magnetic moment per atom, we need the density n of the atoms (the number of atoms per unit volume)

$$n = \frac{\text{atoms}}{\text{volume}} = \frac{\text{mass}}{\text{volume}} \cdot \frac{\text{atoms}}{\text{mass}}$$

$$= \frac{\text{mass}}{\text{volume}} \cdot \frac{\text{atoms/mole}}{\text{mass/mole}}$$

$$= \rho . \frac{N_A}{m}$$

Here ρ is the density of iron, N_A is the Avogadro number, and m is the molar mass of iron. Putting their values, we get

$$n = (7.85 \times 10^3 \text{kg/m}^3) \frac{6.02\times10^{23} \text{ atoms/mole}}{0.0559 \text{ kg/mole}}$$

$$= 8.45 \times 10^{28} \text{ atoms/m}^3.$$

The average magnetic moment of an atom is, therefore,

$$\mu = \frac{M}{n} = \frac{1.11\times10^6 \text{ A/m}}{8.45\times10^{28} \text{ /m}^3} = 1.31 \times 10^{-23} \text{ J/T} = 1.4\mu_B.$$

10.3 Magnetic materials

The absolute permeability μ of a material depends on the magnetizing force, nature and condition of the material. The permeability of an isotropic homogeneous medium is measured by the flux density induced in it by unit magnetizing field. μ is sometimes slightly less and sometimes greater than unity. Also some substances have +ve value of k, while others have –ve k. The differences in the values of k and μ lead us to classify substances into three distinct kinds (1) paramagnetic (2) ferromagnetic, and (3) diamagnetic.

1. Paramagnetism

These materials have a permeability whose value is slightly greater than unity (μ > 1), and become weakly magnetised in the

direction of the magnetizing field. Also k has small positive values. Included in this group are aluminium, chromium, manganese, osmium, palladium, salts of iron, oxygen, platinum. Paramagnetic materials have feeble magnetic properties. For platinum $\mu = 1.00002$ and $k = 1.71 \times 10^{-6}$. Paramagnetic substances move towards stronger parts of a magnetic field, or the lines of force crowd towards such substances. Hence they are attracted by a magnet.

In the absence of an applied field, the atomic dipole moments in a sample of a paramagnetic material initially are randomly oriented in space [Fig. 10.1(a)]. The net dipole moment per unit volume *i.e.*, the magnetisation ($M = \mu/V = \Sigma\mu_i/V$) is zero, because the random directions of the magnetic dipole moments cause the vector

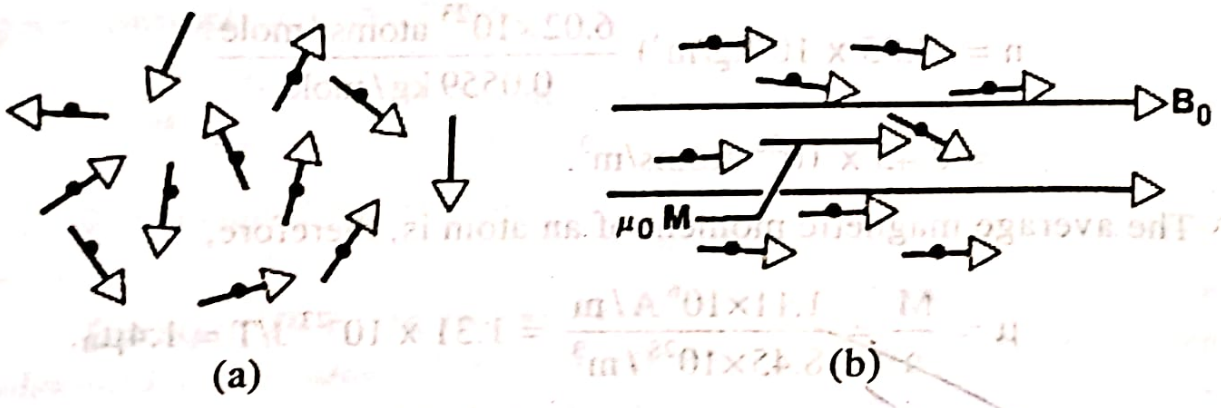


Fig. 10.1

sum to vanish. When an external magnetic field is applied to the material, the resulting torque on the dipoles tends to align them with the field [Fig. 10.1(b)]. The vector sum of the individual dipole moments is no longer zero. The field inside the material now has two components: the applied field (B_0) and the induced field $\mu_0 M$ from the magnetization of the molecules. It may be noted that the two fields are parallel; the dipoles thus enhance the applied field. This is in contrast to the electrical case in which the dipole field opposes the applied field and reduces the total electric field in the material. The ratio between $\mu_0 M$ and B_0, which according to eqn. 10.5 is $k_m - 1$, is small and positive for paramagnetic materials. Relative permeability of some paramagnetic materials at room temperature is given below:

Material	$k_m - 1$
Gd_2O_2	1.2×10^{-2}
$CuCl_2$	3.5×10^{-4}
Chromium	3.3×10^{-4}
Tungsten	6.8×10^{-5}
Aluminium	2.2×10^{-5}
Magnesium	1.2×10^{-5}
Oxygen (1 atom)	1.9×10^{-6}
Air (1 atm)	3.6×10^{-7}

The thermal motion of the atoms which increases with increasing temperature tends to disturb the alignment of the dipoles. As a consequence the magnetization decreases with increasing temperature. The relationship between M and the temperature T was discovered to be an inverse one by Pierre Curie in 1895 and is given by

$$M = C\frac{B_0}{T}$$

which is known as *Curie's law*, the constant C being known as the *Curie constant*.

The magnetization of a particular sample depends on the vector sum of its atomic magnetic dipoles. When all the dipoles are parallel, the magnetization reaches its maximum value. If there are N such dipoles in the volume V, the maximum value of μ is $N\mu_i$, which occurs when all N magnetic dipoles μ_i are parallel. In this case

$$M_{max} = \frac{N}{V}\mu_i$$

When the magnetization reaches this saturation value, any increase in the applied field B_0 will have no further effect on the magnetization. Curie's law, according to which magnetization M should increase linearly with the applied field B_0, is valid only when the magnetization is far from saturation, *i.e.*, when B_0/T is small. Fig. 10.2 shows the measured magnetization M, as a fraction of the maximum value M_{max}, as a function of B_0/T for various temperatures

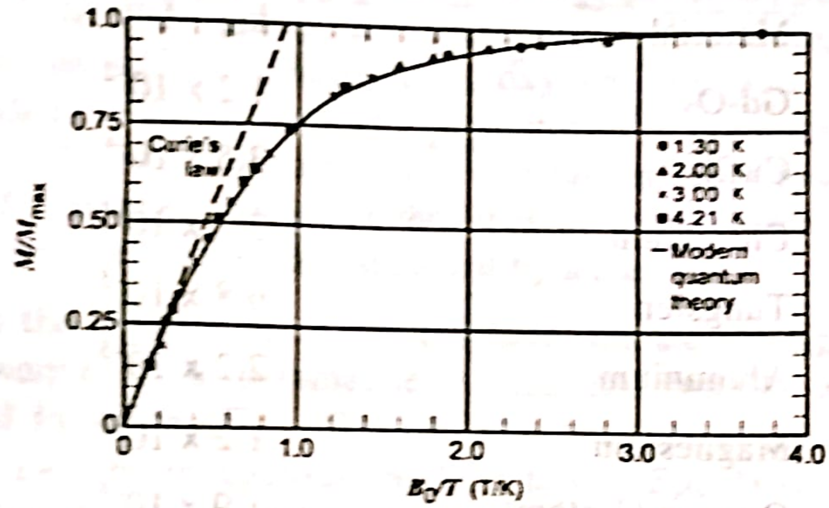


Fig. 10.2

for the paramagnetic salt chrome alum [$CrK(SO_4)_2$, 12 H_2O] in which the chromium ions are responsible for the paramagnetism. As can be seen, Curie's law is valid only for small values of B_0/T, corresponding to either small applied fields or high temperatures.

Paramagnetic Cooling

In 1926 Debye proposed that temperatures very close to absolute zero may be reached by a process known as *adiabatic demagnetization*. Let a sample be magnetized at a constant temperature. The dipoles move into a state of minimum energy in full or partial alignment with the applied field, and in doing so they must give up energy to the surrounding material. This energy flows as heat into the thermal reservoir of the environment. Now the sample is thermally isolated from its environment and the magnetic filed is withdrawn. The thermal motion of the atoms causes the directions of the magnetic dipole moments to become random again. When the dipoles become randomized, the increase in their magnetic energy must be compensated by a corresponding decrease in the internal energy of the system (since heat cannot flow to or from the isolated system in an adiabatic process). The temperature of the sample must therefore decrease. This process of adiabatic demagnetization of atomic magnetic dipoles can be used to achieve temperatures of the order of 0.001K. The demagnetization of the much smaller nuclear magnetic dipoles permits temperatures in the range of 10^{-6} K to be obtained.

2. Diamagnetism

In 1847, Michael Faraday discovered that a specimen of bismuth was *repelled* by a strong magnet. He called such substances diamagnetic. In contrast, paramagnetic substances are always *attracted* by a magnet.

Diamagnetism occurs in materials in which the atoms have no permanent magnetic dipole moment. Such materials might be those having atomic magnetic dipole moments of zero. This might originate from atoms having several electrons with their orbital and spin magnetic moment adding vectorially to zero. When such a material is placed in magnetic field, it acquires *induced* dipole moments. Let us consider the orbiting electrons in an atom to behave like current loops. When an external field is applied, the flux through the loop changes. By Lenz's law, the motion must change in such a way that an induced field opposes this increase in flux. If we were to bring a single atom of a material such as bismuth near the north pole of a magnet, the field (which points away from the pole) tends to increase the flux through the current loop that represents the circulating electron. According to Lenz's law, there must be an induced field pointing in the opposite direction (toward the pole). The induced north pole is on the side of the loop toward the magnet, and the two north poles repel one another.

This effect occurs no matter what the sense of rotation of the original orbit, so the magnetization in a diamagnetic material opposes the applied field. Thus the diamagnetic materials, when placed in a magnetic field, have a tendency to move away from the field. For diamagnetic materials $\mu < 1$, k is constant and has –ve value. The ratio of the magnetization contribution to the field B_0, given by $k_m - 1$, amounts to -10^{-6} to -10^{-5} for typical diamagnetic materials, *e.g.*, bismuth, antimony, Zn, Ag, Cu, Sb, Au, Pb, water, alcohol, hydrogen and inert gases.

The relative permeability of some diamagnetic substance at room temperature is given below:

Substance	$k_m - 1$
Mercury	-3.2×10^{-5}
Silver	-2.6×10^{-5}
Bismuth	-1.7×10^{-5}
Ethyl alcohol	-1.3×10^{-5}
Copper	-9.7×10^{-6}
Carbon dioxide (1atm)	-1.1×10^{-8}
Nitrogen (1 atm)	-5.4×10^{-9}

3. Ferromagnetism

Like paramagnetic materials, ferromagnetic are those in which the atoms have permanent magnetic dipole moments. What distinguishes ferromagnetic materials form paramagnetic materials *is that in* ferromagnetic materials there is a strong interaction between neighbouring atomic dipole moments that keeps them aligned even when the external magnetic field is removed. Whether or not this occurs depends on the strength of the atomic dipoles and, because the dipole field changes with distance, on the separation between the atoms of the material. Because the separation depends on the temperature, some materials like iron, cobalt, steel, nickel, show ferromagnetism at room temperatures, while elements of rare earth such as gadolinium or dysprosium show ferromagnetism only at temperatures much below room temperature.

Ferromagnetic materials have an abnormally high +ve value of k. The total magnetic field **B** inside a ferromagnet may be 10^3 or 10^4 times the applied field B_0. As in paramagnetic materials, the magnetization produced in ferromagnetic materials is not proportional to the magnetizing force; hence k and μ vary with the magnetizing force considerably. Also k varies with temperature. As the effectiveness of the coupling between neighbouring atoms that causes ferromagnetism by increasing the temperature of the substance, its k varies inversely as the absolute temperature. This is called **Curie law**, and is expressed as kT = constant, where k is the susceptibility and T is the absolute temperature. Thus k decreases steadily with rise of temperature, until a *critical temperature* called the **curie temperature** or **curie point** is reached. At this temperature ferromagnetism disappears and the substance becomes paramagnetic. The susceptibility of a ferromagnetic substance above its curie point is inversely proportional to the amount its temperature is above the curie point. This is called the **Curie Weiss law**. The curie point is about 100°C for cobalt, 400°C for nickel and 770°C for iron. Above 770°C, iron is paramagnetic. The curie temperature of gadolinium metal is 16°C; at room temperature gadolinium is paramagnetic, while at temperature below 16°C, gadolinium becomes ferromagnetic.

We shall now discuss some characteristics of ferromagnetic substances.

10.4 Retentivity and Coercivity

If pieces of soft iron and steel of the same dimensions are subjected to the same magnetizing force, it is found that the two metals do not retain the same degree of magnetization when the force is withdrawn. If the metals are not subjected to any disturbing influence, the soft iron may retain as much as 90% of its original magnetization, while steel retains much less than this amount. The power of retaining magnetization after the removal of the magnetizing force is called *retentivity.* The residual magnetism or for that matter the whole of the original magnetization may be brought to zero by subjecting the metals to a reverse magnetizing force. The degree of this reverse magnetizing force which is necessary to deprive a metal of the whole of its original magnetization, is called the *coercive force* – and the capacity of retaining magnetization inspite of any subsequent treatment is called *coercivity.* It has been found that although soft iron has much greater retentivity than steel, steel has far greater coercivity than soft iron.

10.5 Cycle of magnetization: Hysteresis

Let us insert a ferromagnetic material such as iron into a solenoid. We assume that the current is initially zero and the iron is unmagnetized, so that initially B_0 (or H) *i.e.*, the magnetizing field, M, the intensity of magnetization and the magnetic induction B are zero. As H is increased by increasing the current in the solenoid, the magnetization increases rapidly as indicated by the portion OAC in Fig. 10.3. If the magnetizing force is increased beyond the point C, M remains constant and the substance is said to be *saturated.* B, however, does not remain constant after the saturation point has been reached, but increases very slightly with increase in the magnetizing field.

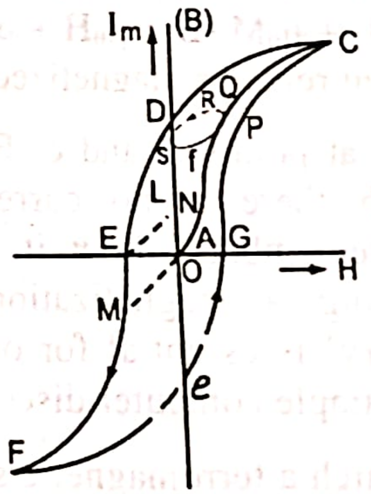


Fig. 10.3

If now, when the point C is reached, H is gradually decreased and the corresponding value of the intensity of magnetization is noted, it will be found that the curve does not coincide with that obtained with increasing value of H, but takes the form CD. Thus when the magnetizing force is zero again (H = O), the intensity of magnetization I_m (or the induction B) instead of being zero has a value = OD. This value of intensity of magnetization for which H = O, is called *residual magnetism* or *retentivity* or *remanence*.

If the direction of H is reversed, the curve DEF is obtained, the point F being the saturated magnetization in the opposite direction. If the current is now returned to zero we find that the sample retains a permanent magnetization at point e (Fig. 10.3) If the current is again increased to return to the saturated magnetization in the original direction (point C), the curve e GPC is obtained. Thus in all cases I_m and hence the magnetization (similarly induction B) appears to lag behind the magnetizing force. This lagging of intensity of magnetization (or induction B) behind the magnetizing force is called *hysteresis*, and the closed loop obtained (CDEFeGC) is called *hysteresis loop*. The cycle of operation is called *hysteresis cycle*. The magnetizing force represented by OE or OG, represents the force required to deprive the bar of its residual magnetism (M = 0) and it, therefore, represents the *coercive force* for the material.

The shape of the hysteresis loop, when B is plotted against H, is similar to the one between M and H, although the areas of the loops are different. The loop is then also referred to as B – H curve. The intercept on the B-axis when H = 0 is called remanent induction. It must be noted that the value of the applied field to make B = 0 is not coercivity; because $B = B_0 + \mu_0 M$ ($B = \mu_0 H + \mu_0 M$) and when B = 0, H = M so that the specimen remains magnetized *i.e.*, M is not zero.

It may be noted that at points D and e (Fig. 10.3), the iron is magnetized, even though there is no current in the solenoid. Furthermore, the iron "remembers" how it became magnetized, negative current producing a magnetization different from a positive one. This 'memory' is essential for operation of magnetic storage information for example computer discs and cassette tapes.

The mechanism by which a ferromagnetic substance approaches saturation value is different from that of paramagnetic substance. In

case of a paramagnetic material the saturation value is approached by rotation of the individual dipole moments so as to align them with the applied field. A ferromagnetic material such as iron, on the other hand, is composed of a large number of microscopic crystals. Within each crystal are *magnetic domains*, regions of roughly 0.01 mm in size in which the coupling of atomic magnetic dipoles produces essentially perfect alignment of all the atoms. There are many domains, each with its dipoles pointing in a different direction The net result of adding these dipole moments in an unmagnetized ferromagnet gives a magnetization of zero.

When the ferromagnet is placed in an external field, two things may happen : (i) dipoles outside the walls of domains that are aligned with the field can rotate into alignment, in effect allowing such domains to grow at the expense of neighbouring domains; (ii) and the dipoles of the non-aligned domains may swing entirely into alignment with the applied field. In either case, there are now more dipoles which are aligned with the field. The material therefore has a large magnetization. When the external field is removed, the domain walls do not fall back completely to their former positions. The material retains a magnetization in the direction of the applied field.

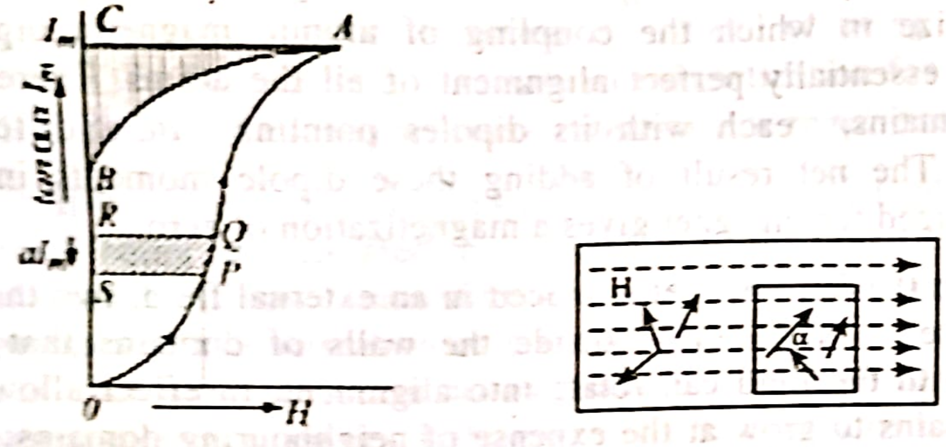
10.6 Energy dissipation due to hysteresis –Hysteresis loss

The act of taking a ferro-magnetic substance through a cycle of magnetization involves the expenditure of energy. This is because work has to be done to increase the induction and the greater the induction the greater the work that has to be done to increase the induction by a given amount. Thus, during a cycle, a greater magnetizing force has to be used to obtain a given induction while magnetizing a substance, than that which corresponds to the same induction when the magnetizing force is decreasing. The energy required to magnetize a specimen cannot be recovered as the magnetism does not become zero on removal of the magnetizing field; a field in the reveres direction has to be applied before M (or B) can be brought to zero. Thus there is a loss of energy when a magnet is taken through a cycle and it appears as heat in the specimen.

Calculation of energy dissipated

Fig. 10.4 shows a part OAB of the hysteresis loop drawn between the magnetizing field H, which is proportional to the current i (the

magnetizing current) and tan α, which is proportional to the intensity of magnetization M. A similar curve will be obtained between H (proportional to i) and α which is proportional to B (the magnetic induction). However, we shall use the relation between H and M.



(a) (b)

Fig. 10.4

Let μ_i = magnetic moment of one of the molecular magnets of the specimen [Fig. 10.4(b)] α = angle between the magnetic axis of the molecular magnet and the direction of the magnetizing field H.

$$\text{Component of } \mu_i \text{ along } H = \mu_i \cos\alpha$$

Sum of magnetic moment of all the molecular magnets in *unit volume* of the specimen in the direction of H

$$= \Sigma\, \mu_i \cos\alpha = M \text{ (intensity of magnetization).}$$

If M is increased by an amount dM, then μ_i increases.

$$\therefore\ d\Sigma\mu_i\cos\alpha = \mu_0 dM; \qquad \text{or} \qquad -\Sigma\mu_i \sin\alpha\, d\alpha = \mu_0 dM$$

The couple acting on each molecular magnet

$$= \mu_i H \sin\alpha$$

So, the work done on all molecular magnets in unit volume in changing their direction through $d\alpha$

$$= \Sigma\mu_i H \sin\alpha\, d\alpha$$

$$= \mu_0 H dM$$

= area SPQR ($d\alpha$ is the reduction in the angle α)

Hence the total work done in completely magnetizing the specimen = sum of such small areas into which the figure OPACBO may be divided = area OACO (i)

On demagnetization, amount of work recovered

= Area ACB (Vertical shading) (ii)

Therefore loss in energy

W = area OACO – area ACB

= area OABO

= area of the M – H loop.

$$= \oint \mu_0 \, HdM \qquad (10.6)$$

A similar loss of energy will occur in the lower part of the hysteresis loop. Thus, the net work done in taking a specimen *completely through hysteresis cycle* equals the total area enclosed by the H-M loop.

Eqn. 10.6 is referred to as Warburg's law

The symbol $\oint$ indicates that the integration is carried out round a complete cycle.

The loss of energy is converted into heat

$$H = \frac{\text{area of M - H loop}}{4.2}$$

If, however B – H loop, instead of M – H loop is drawn, it can be shown that the area enclosed by the B – H loop is 4π times the work done for M – H loop.

The loss of energy for a complete cycle of B – H loop = area of the loop = $\oint H.dB$ where B is in amp/m^2, H in ampere turns and loss of energy in joules.

Thus, the energy wasted in taking a ferromagnetic substance through a complete cycle of magnetization is proportional to the area of the hysteresis loop and is converted into heat within the substance. A rough measure of this is given by the product of remanence and coercivity.

If a volume V of iron is carried through f cycles per second, the hysteresis loss is

$$W = \frac{Vf\ (\text{area of hysteresis loop})}{4\pi}\ \text{watts.} \tag{10.7}$$

It should be mentioned that the area enclosed by the hysteresis loop for a given material depends on the maximum value of B (or M) attained. The area enclosed is greatest when the material is taken to saturation.

Practical importance of hysteresis curves

Iron can be magnetized to a great extent by means of weaker fields than cobalt and nickel; as such it is extensively used where strong magnetization is required. Pure iron is one of the most important magnetically soft materials; but since it is too soft mechanically and rusts too easily, it is not suitable for industrial purpose. But the magnetic properties of iron can be varied very considerably by alloying it with other metals, and *the hysteresis curves for them afford valuable information required for the selection of materials for different industrial purposes.*

For commercial use, two types of magnetic materials are required, (a) soft and (b) hard materials.

For **transformer and armature cores** a soft material is used when it is to be subjected to alternating field, *e.g.* in the cores of transformers, telephone diaphragms and in armature cores of motors and dynamos. In these devices, the material is carried through a cycle of magnetization many times per second by the alternating field. The area of M–H or B–H loop must, therefore, be small; otherwise good deal of energy will be wasted as heat, and it may be difficult to avoid overheating. The materials should also have high resistivity to reduce eddy current losses which is desirable for the above devices. Also they should have high permeability at a small value of the magnetizing field. For the core of an electromagnet, the material should be such that a small magnetizing field can produce a large induction. The hysteresis loops of such substances are very thin. Silicon iron (4% Silicon), rather than soft iron, which satisfies the above requirement is a suitable material for these devices.

For the true reproduction of waveforms, the relation between B and H should be linear over a wide range of magnetization for the core materials of communication transformers. This means that the permeability should be high and constant over a wide range *i.e.*, the hysteresis loop should have as small an area as possible having very slight curvature.

Thus in transformers and other appliances where alternating current is used, in which the material is taken round a hysteresis cycle, alloys of iron, rather than pure iron, are used.

The chief requirement for an *electromagnet* and a *relay* is a large intensity of magnetization for a small current, A low coercive force is desirable so that it can lose its magnetization when the current is switched off. Again steel iron (iron containing 4% silicon) and an alloy of 50% iron and 50% cobalt are used in the design of electromagnet.

For a material to be used as a permanent magnet, hysteresis loss and resistivity are not important as the material is never taken through a cycle of magnetization. Even the permeability may be small as the materials may be magnetized strongly initially. The main requirement is that they should have great remanence, and should have high coercivity so that they are difficult to demagnetize. Hence their M–H or B–H curves are broad. The presence of chromium or tungsten, particularly the latter, increases the circuitry to a great extent. Hence tungsten steel is largely used for permanent magnets.

Example 10.2 *A paramagnetic substance is composed of atoms with a magnetic dipole moment of $3.3\mu_B$. It is placed in a magnetic field of strength 5.2 T. To what temperature must the substance be cooled so that the magnetic energy of each atom would be as large as the mean translational kinetic energy per atom?*

Soln.

The magnetic energy of a dipole in an external field is U = − **μ.B**, and the mean translational kinetic energy per atom is $\frac{3}{2}$kT (see kinetic theory of gases). These are equal in magnitude when the temperature is

$$T = \frac{\mu B}{(3/2)K} = \frac{(3.3)\left(9.27\times10^{-24}\,J/T\right)(5.2T)}{(1.5)\left(1.38\times10^{-23}\,J/K\right)}$$

$$= 7.7\text{ K.}$$

***Example 10.3** An iron rod of density 7.7 and sp. heat 0.1 is subjected to magnetization at 50 cycles per second. If the area enclosed by I_m–H curve for the specimen is 5×10^4 erg, find the rise in temperature per minute ($J = 4.2 \times 10^7$ erg/cal).*

Soln.

Work done per c.c. per minute = area of I_m–H loop

$= 5 \times 10^4$ erg.

Total work done per c.c. per minute $= 5 \times 10^4 \times 50 \times 60$ erg.

$$\text{Heat developed is } H = \frac{W}{J} = \frac{5\times10^4\times50\times60}{4.2\times10^7} = \frac{25}{7}\text{ cals.}$$

Since m = 7.7 gm; sp.ht. = 0.1, and $H = \dfrac{25}{7}$

we have from $ms\theta = H$

$$7.7 \times 0.1 \times \theta = \frac{25}{7};\text{ or } \theta = \frac{25}{7\times7.7\times0.1} = 4.64°\text{C.}$$

$$= 277.64\text{K.}$$

EXERCISES

1. Discuss briefly the distinguishing features of diamagnetic, paramagnetic and ferromagnetic substances.

2. Define magnetic force, magnetic induction, intensity of magnetization, permeability and susceptibility. Find the relation between these quantities.

3. What is meant by hysteresis? Describe a method by which you would obtain the hysteresis curve for a ferromagnetic substance. How does the study of hysteresis help us in the choice of magnetic materials?

4. What is meant by hysteresis, residual magnetism, coercive force and permeability? Describe a method for obtaining hysteresis loop.

5. Describe the formation and significance of the hysteresis loop and explain the terms hysteresis loop and explain the terms hysteresis, cycle of magnetisation. Show that the energy dissipated in unit volume of iron during a complete hysteresis cycle is $\frac{1}{4\pi}$ times the area enclosed by B–H curve.

6. Prove that the area of the B–H cycle is 4π times the energy dissipated per cc. of the metal during each magnetic cycle.

CHAPTER - XI

SPECIAL THEORY OF RELATIVITY

11.1 The search for a frame of reference – the ether

The problem of absolute motion of bodies, *i.e.*, motion with respect to a fixed frame of reference and the interpretation of various experimentally observed facts in terms of absolute motion, has always been the burning question in physics. Newton was forced to admit that the absolute motion of a uniformly moving system cannot be detected by mechanical experiments conducted within the system. But then, around the middle of the nineteenth century, James Clerk Maxwell demonstrated that electricity and light are related phenomena. He further demonstrated that the velocity of electromagnetic waves was the same as the velocity of light. He could derive many other properties of light, and it was soon accepted that light is an electromagnetic wave motion.

Every wave motion has something that waves. Sound waves have air and water waves have water. It was, therefore, argued that light waves must involve the waving of something even in free space. No one knows what it was, but the name *luminiferous ether* was given to it.

To account for the propagation of light through vacuum as well as material medium, the high velocity of light, transverse nature of light, ether was assigned many properties, often contradictory in nature. But because of the success of the wave theory of light in explaining interference, diffraction, polarization and other phenomena of light, the ether gradually came to be accepted as a physical reality.

The ether was assumed to be stationary and this gave rise to new possibility. It was inferred that it should be possible to determine the absolute velocity of a body, *i.e.*, velocity relative to stationary ether. Since all speculations about the ether stem from its properties as a medium for carrying light, an optical experiment is indicated. It was thought that the measurement of absolute velocity should depend on the effect of motion on some phenomena involving light. Thus if light and the earth both move with definite velocities relative to the stationary ether, it should be possible to

devise an experiment, involving the propagation of light, from which the absolute velocity of the earth could be deduced. Of the many such experiments proposed, the theoretical basis of the most famous one – the *Michelson-Morley experiment*, can be easily derived.

The Michelson – Morley Experiment

It was A. A. Michelson who invented the optical interferometer whose remarkable sensitivity made the experiment possible. Michelson first performed the experiment in 1881. Then in 1887, in collaboration with E. W. Morley, he carried out the more precise version of the investigation that was destined to lay the experimental foundations of the theory of relativity. For this invention of the interferometer and his many optical experiments, Michelson was awarded the Nobel Prize in Physics in 1907, the first American to be so honoured.

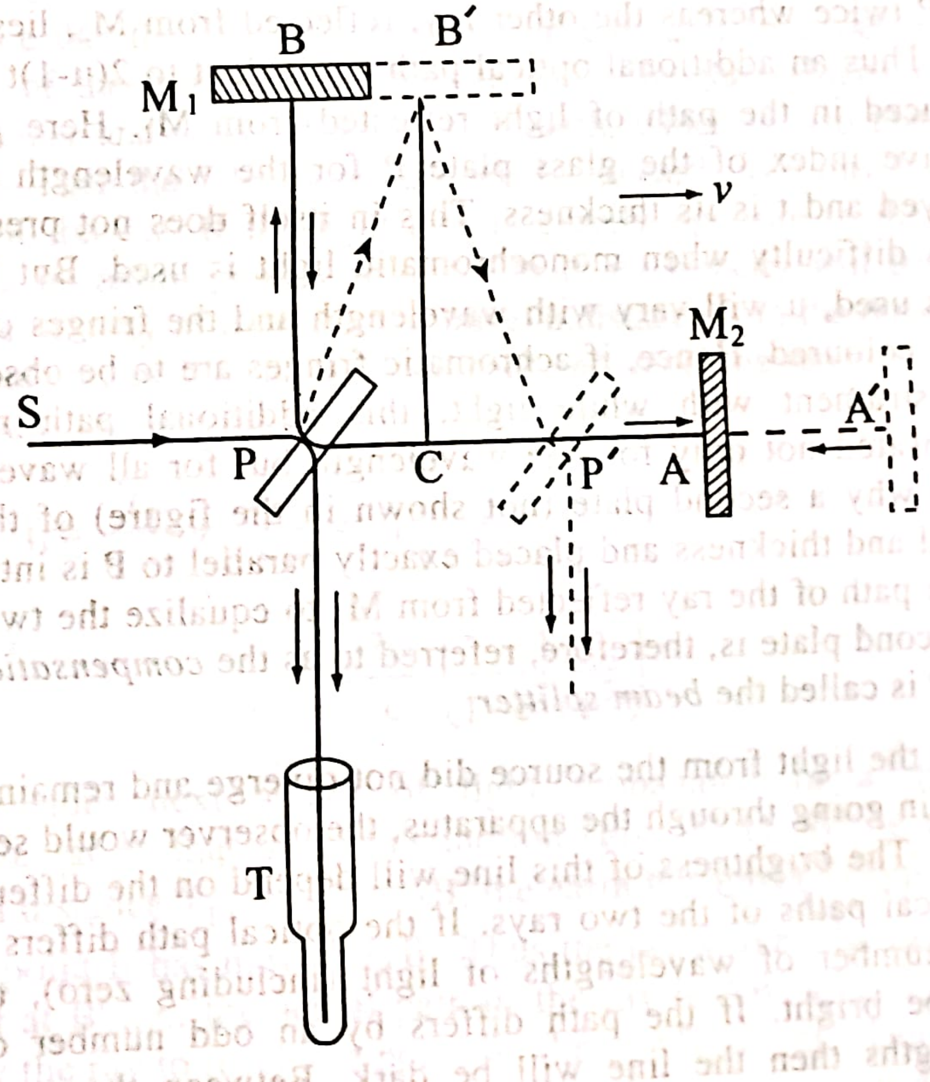


Fig. 11.1

Let us now describe the Michelson interferometer. Light from an extended monochromatic source S falls on a glass plate P (Fig. 11.1) placed at 45° to the beam. P is half-silvered on its right side and this surface reflects half of the light up towards M_1 while the other half is transmitted and goes towards M_2. The reflected portion travels in a direction at right angles to that of the initial beam and falls normally at B on M_1. The transmitted portion travels along the direction of the initial beam and falls normally at A on M_2. Both M_1 and M_2 are full silvered front surface mirrors that reflect their beams back towards P. The beam from M_1 is partly reflected at P and the remainder goes on through to the telescope T. A portion of the beam from M_2 is reflected at P to the telescope T and the rest goes through the glass plate and is lost. The two rays thus returned to P by M_1 and M_2 interfere on their final journey towards T.

It may be noted that the ray reflected from M_1 crosses the glass plate P twice whereas the other ray, reflected from M_2, lies wholly in air. Thus an additional optical path equivalent to $2(\mu-1)t$ in air is introduced in the path of light reflected from M_1. Here μ is the refractive index of the glass plate P for the wavelength of light employed and t is its thickness. This in itself does not present any serious difficulty when monochromatic light is used. But if white light is used, μ will vary with wavelength and the fringes observed will be coloured. Hence, if achromatic fringes are to be observed in this instrument with white light, this additional path must be compensated not only for one wavelength but for all wavelengths. That is why a second plate (not shown in the figure) of the same material and thickness and placed exactly parallel to P is introduced into the path of the ray reflected from M_2 to equalize the two paths. This second plate is, therefore, referred to as the *compensation plate* while P is called the *beam splitter*.

If the light from the source did not diverge and remained very narrow in going through the apparatus, the observer would see a line of light. The brightness of this line will depend on the difference in the optical paths of the two rays. If the optical path differs by any whole number of wavelengths of light (including zero), the line would be bright. If the path differs by an odd number of half-wavelengths then the line will be dark. Between these extremes ... brightness gradation would be observed. In practice, the light

does diverge in the apparatus and a great many slightly different paths are being traversed simultaneously. Consequently, the observer does not see one line but a multiplicity of lines. The loci of points where the paths differ by whole wavelengths are bright while the loci of the points where the paths differ by an odd number of half-wavelengths are dark. Thus, as one path length is varied, the observer sees fringes move across the field, rather than a single line becoming lighter and darker. It is rather fortunate that the optical system works as it does, since it is easier for the eye to detect differences in position than differences in intensity.

Let l_1 be the distance of M_2 from P and l_2 that of M_1 from P. In the actual experiment, the two arms of the interferometer were made nearly equal, *i.e.*, $l_1 = l_2 = l$ (say). If the experimental arrangement were at rest in ether, the two rays would take the same time to return to P. But in the actual experiment the whole arrangement is moving with the earth. Let us suppose that the direction of motion of the earth coincides with the direction of the initial beam of light. Let c be the velocity of light through ether and v be the velocity of earth, which is also that of the apparatus. Then the velocity of light relative to earth, as it travels from P to M_2 will be c-v. If the light travels exactly opposite to the earth's motion in its return journey from M_2 to P, its velocity will be c + v. Therefore, if a ray of light were to travel a distance l_1 in the direction of earth's motion, and then be reflected back and travel the same distance in the opposite direction the time required would be

$$t_1 = \frac{\ell_1}{c-v} + \frac{\ell_1}{c+v} = \ell_1\left(\frac{2c}{c^2 - v^2}\right)$$

$$= \frac{2\ell_1}{c}\left(\frac{1}{1 - c^2/v^2}\right)$$

Consider next a ray of light propagated in a direction perpendicular to that of earth's motion (Fig. 11.1). While this ray travels a distance ℓ_1, equal to AB, the earth has moved from A to A' and the point B has moved to B'. Thus the ray strikes the mirror not at B but at B' *i.e.*, the actual path of the ray is AB'. If t is the time taken by the ray to reach the point B', then AB' = ct. During the same interval, A has moved to A' with the velocity v, so that AA' = vt. In the right angled triangle AA'B, we have

$$c^2t^2 = l_2^2 + v^2t^2$$

$$\text{or,}\quad t^2(c^2 - v^2) = l_2^2$$

$$\text{or,}\quad t = \frac{l_2}{\sqrt{c^2 - v^2}}$$

The same length of time will be needed for the return trip, *i.e.*, to A". Hence, the time required to travel a distance l, and back in a direction perpendicular to that of the earth's motion is

$$\text{or,}\quad t_2 = 2t = \frac{2l_2}{\sqrt{c^2 - v^2}} = \frac{2l_2}{c} \cdot \frac{1}{\sqrt{1 - v^2/c^2}}$$

The calculation of t_2 is made on the ether frame, that of t_1 in the frame of the apparatus. Because time is absolute in classical physics, this is perfectly acceptable classically. It may be noted that both effects are second-order ones ($v^2/c^2 \simeq 10^{-8}$) and are in the same direction (they increase the transit time over the case v = 0). Hence, the difference in transit times of the parallel and perpendicular rays

$$\Delta t = t_2 - t_1 = \frac{2}{c}\left[\frac{l_2}{\sqrt{1 - v^2/c^2}} - \frac{l_1}{1 - v^2/c^2}\right]$$

Suppose that the instrument is rotated through 90°, thereby making l_1 the perpendicular length and l_2 the length in the parallel direction. If the corresponding times are now designated by primes, the same analysis as above gives the transit-time difference as

$$\Delta t' = t'_2 - t'_1 = \frac{2}{c}\left[\frac{l_2}{1 - v^2/c^2} - \frac{l_1}{\sqrt{1 - v^2/c^2}}\right]$$

Hence, *the rotation changes the differences* by

$$\Delta t' - \Delta t = \frac{2}{c}\left[\frac{l_2 + l_1}{1 - v^2/c^2} - \frac{l_2 + l_1}{\sqrt{1 - v^2/c^2}}\right]$$

Using the binomial expansion and neglecting terms higher than the second-order (v^2/c^2), we may write

$$\frac{1}{1-v^2/c^2} = 1 + \frac{v^2}{c^2} + \ldots\ldots$$

and $$\frac{1}{\sqrt{1-v^2/c^2}} = 1 + \frac{1}{2}\frac{v^2}{c^2} + \ldots\ldots$$

Hence,

$$\Delta t' - \Delta t \cong \frac{2}{c}(l_1 + l_2)\left[1 + \frac{v^2}{c^2} - 1 - \frac{1}{2}\frac{v^2}{c^2}\right]$$

$$= \left(\frac{l_1 + l_2}{c}\right)\frac{v^2}{c^2}$$

Therefore, the rotation of the apparatus through 90° should cause a shift in the fringe pattern since a transit-time difference ($\Delta t'$-Δt) will introduce an optical path difference ($\Delta t' - \Delta t$).c between the two rays.

If the optical path difference between the rays changes by one-wavelength, for example, there will be a shift of one fringe across the cross-wires of the viewing telescope. Let ΔN represent the number of fringes moving past the cross-wires as the pattern shift due to the path difference ($\Delta t' - \Delta t$).c. Then if light of wavelength λ is used, we have

$$(\Delta t' - \Delta t).c = \Delta N.\lambda$$

or, $$\Delta N = \frac{(\Delta t' - \Delta t).c}{\lambda}$$

$$= \frac{(l_1 + l_2)}{c} \cdot \frac{v^2}{c^2} \cdot \frac{c}{\lambda}$$

$$= \frac{(l_1 + l_2)}{\lambda} \cdot \frac{v^2}{c^2} \qquad (11.1)$$

In the actual experiment, the two arms of the interferometer were of (nearly) equal length, that is $l_1 = l_2 = l$. Michelson and Morley were able to obtain an optical path length, $l_1 + l_2$, of about 22 m. The wavelength of light used, $\lambda = 5.9 \times 10^{-7}$ m, $v = 3 \times 10^4$ m/sec and $c = 3 \times 10^8$ m/sec. We, therefore, obtain from eqn. 11.1

$$\Delta N = \frac{22}{5.9\times10^{-7}} \times \frac{(3\times10^4)^2}{(3\times10^8)^2}$$

$$= 0.37 \text{ fringes,}$$

or a shift of about four-tenths of a fringe.

Michelson and Morley estimated that they could detect a shift of one-hundredth of a fringe. Sensitivity to spare !

In the actual carrying out of the experiment, in order to *secure constancy of temperature as well as avoid vibration effects*, the dimensions of the apparatus were reduced by making the arms PA and PB small, but to make the light path as long as possible, mirrors were arranged on the slab to reflect the beams back and forth a great number of times before returning to the plate P. The whole apparatus was mounted on a massive slab of sand stone for stability and floated in a pool of mercury so that it could be rotated smoothly about a central pin.

Measurements were made over an extended period of time under a continuous slow rotation of the apparatus making one revolution in about six minutes. Observations were made day and night (as the earth spins about its axis) and during all seasons of the year (as the earth rotates about the sun), but the expected fringe shift was not observed. The result obtained was surprisingly unexpected. The actual shift observed by them was not only *not consistent* but also certainly less than one-twentieth of what was expected. Indeed, the experimental conclusion was that *there was no fringe shift.*

The failure to detect any fringe shift ($\Delta N = 0$) was such a blow to the ether hypothesis, that the experiment was repeated by many workers over a 50 years period. The null result ($\Delta N = 0$) was amply confirmed in all these experiments. In 1958, J.P. Cedarholm, C.H. Townes et al repeated the experiment using micro-waves. The improvement in the precision of the experiment was 50 times over the best experiment of the Michelson-Morley type. The experiment showed that if there is an ether and the earth is moving through it, then the earth's speed with respect to ether would have to be less than 1/1000 of the earth's orbital speed. The null result was again confirmed.

It should be noted that the Michelson - Morley experiment depends essentially on the 90° rotation of the interferometer, that is, on interchanging the roles of l_1 and l_2, as the apparatus moves with a speed v through the ether. In predicting an expected fringe shift, v was taken to be the earth's velocity with respect to an ether fixed with the sun. However, the solar system itself might be in motion with respect to the hypothetical ether. Then the negative result of the experiment might just have been due to the fact that at the time of the experiment the earth might have had no resultant component of velocity parallel to the surface of the earth due to the motion of the solar system as a whole. It might happen for the orbital and solar motion to be in opposite senses and cancel each other at a given time of year. But then six months later the earth's orbital velocity will be reversed and in consequence its velocity through ether would be twice its orbital velocity. As mentioned earlier the experiment was repeated by Michelson and Morley and many others at different times of the year when the directions of the orbital velocity are different. The experiment has since been repeated below the earth's surface as well as at high altitudes at different times of the year and, lately, with a highly monochromatic light from a laser, but the result, in all cases has been a negative one. Other experiments to detect the presence of ether have similarly failed. All these experiments gave the same negative result.

One way to interpret the negative result of Michelson - Morley experiment is to conclude simply that the measured speed of light that is, c, is the same for all directions in every inertial system. Thus light traveled with the same speed in the two directions in Michelson-Morley experiment and if the two arms of the interferometer are equal, this will lead to $\Delta N = 0$. However, such a conclusion, seemed to be too drastic philosophically at the time as it was incompatible with Galilean (velocity) transformation. The simplest interpretation of the negative result seemed to be that v, the velocity of earth relative to ether, is zero; or there is no existence of a unique inertial system, that is, the ether.

Explanation of the negative result

Few experimental failures have been more stimulating than the failure of Michelson-Morley experiment to detect the existence of ether. The theoretical value of the fringe shift was derived by simple and straight-forward application of some of the fundamental ideas of

classical physics. This fact presented a challenge to explain the negative result of the Michelson-Morley experiment.

The ether-drag hypothesis: The simplest explanation was to propose that the moving earth dragged the ether with it, in which case there would be no relative motion between the two (the velocity of the earth relative to ether should be zero). This automatically gives a null result in the Michelson-Morley experiment. However, the ether-drag theory was unattractive because every moving body would presumably carry its own ether with it, and the uniqueness of the stationary ether as the medium for the propagation of light would disappear.

There were also experimental objections to the idea that the earth carries the ether along with it, for, it can be shown that if the earth drags the ether with it then the velocity of light will depend on the velocity of the source of light. All observations on the velocity of light contradicts this conclusion. The velocity of light was always found to be the same regardless of whether or not the emitting source moves or how it moves. The assumption that the earth carries the ether along with it led, therefore, to a serious contradiction with experiment and had to be rejected.

The Lorentz - Fitzgerald Contraction Hypothesis

L.F. Fitzgerald, an Irish physicist, in an attempt to explain the negative result of Michelson-Morley experiment while preserving the idea of the stationary ether, introduced a hypothesis. According to this hypothesis, there might be an interaction between the ether and objects moving relative to it, such that the objects contract by a factor $\sqrt{1-v^2/c^2}$ in all its dimensions parallel to the relative velocity, while there is no such contraction in a direction perpendicular to the relative velocity. Thus the arm of the interferometer that is parallel to the direction of the relative velocity contracts in length from l to $l\sqrt{1-v^2/c^2}$, but the length of the perpendicular arm remains equal to l. Therefore, *the light does not travel as far in the parallel direction as in the perpendicular direction*, and the difference is such as to make theoretical value of the shift of the interference pattern equal to zero. The contraction could never be measured because any rule used to measure it would contract also.

Fitzgerald hypothesis seems absurd at first sight, because the contraction was not the result of any forces acting on the body, but it depends only on the fact that the body is in motion. It was contended by Fitzgerald that because of the electrical nature of the matter, the size and shape of a solid body, depending on the interaction of electrical forces between the molecules, may be altered by the motion of the body through ether perfectly at rest. Fitzgerald's hypothesis was, therefore, purely ad-hoc; its main object was to preserve the concept of a stationary ether. Lorentz, who was at the same time trying to develop a consistent electron theory of matter found Fitzgerald's hypothesis useful and he incorporated it into his theories, so that the hypothesis is now known as the *Lorentz-Fitzgerald contraction hypothesis*.

The mass of an electron, under certain assumptions, could be expressed by the formula

$$m_0 = \frac{2e^2}{3r_0c^2}$$

where r_0 is the radius of the electron when it is at rest, and e is its charge. Lorentz assumed that a moving electron contracts in the direction of the motion by the Fitzgerald factor $\sqrt{1-v^2/c^2}$, so that the radius becomes $r_0\sqrt{1-v^2/c^2}$ when the electron moves with a velocity v. Then if the mass of the moving electron is m, then

$$m = \frac{2e^2}{3r_0\sqrt{1-v^2/c^2}}$$

$$= \frac{m_0}{\sqrt{1-v^2/c^2}} \qquad (11.2)$$

where m_0 is the mass of the electron when it is at rest.

Thus, according to eqn. 11.2 the mass of an electron should increase with its velocity, especially when its velocity reaches a significant fraction of the velocity of light. Early experiments on the value of e/m for β-rays indicated that the value of e/m decreases

with increasing velocity of the rays. Since there is no reason to suppose that the electronic charge varies with velocity, it was concluded that the mass of an electron does indeed increase with its velocity. Thus, although Lorentz-Fitzgerald contraction hypothesis seemed physically unreasonable, it led to theoretical results consistent with experiment. But the hypothesis lacked a sound theoretical basis. It was at the most an ad-hoc hypothesis proposed to get out of a particular difficulty (explanation of the negative result of Michelson-Morley experiment) without any positive experimental confirmation. Lord Rayleigh, in fact, pointed out that such deformation of matter of purely electromagnetic origin should give rise to double refraction. Lord Rayleigh himself and then Brace carefully sought for this effect. But this effect as well as other results predicted on the basis of this contraction hypothesis (electron theory of matter) were never observed.

11.2 Einstein's concept of relativity

The negative result of Michelson-Morley experiment had two consequences. **First**, it rendered untenable the ether hypothesis demonstrating that the ether has no measurable properties – an ignominious end for what had once been a respected idea. **Second**, it suggested a new physical principle; the speed of light in free space is same everywhere, regardless of any motion of the source or the observer. The difficulties and anomalies encountered in the different solutions of the problems of absolute motion through ether should have suggested to the scientists that the very concept on which the question was framed might after all be wrong. As a matter of fact the French mathematical physicist Henri Poincare came very close when he stated that '.......... *the laws of physical phenomena (are) the same, whether for an observer fixed or for an observer carried along in a uniform movement of translation, so that we have not and could not have any means of discerning whether or not we are carried along in such a motion*'.

This simply means that if we are drifting with uniform speed in a spaceship, with all the windows closed, we shall not be able to say, with the help of any experiments we might choose to perform, whether we are at rest or in motion. If, however, we look out of a window, and see another spaceship change its position relative to us,

we can merely say that we are in motion relative to the second spaceship but not whether we or the second spaceship are actually in motion. Should we be isolated in a universe, there would be no way in which we could determine whether we are in motion or not, because without a frame of reference, the concept of motion has no meaning.

How near, indeed, had Poincare thus come to expounding the theory of relativity and yet how far he actually was from it. For instead of grasping the implications of the negative result of all ether drift experiments, and expounding a new theory on its basis, discarding old concepts of space and time, he devoted himself to somehow trying to save the old classical theory by suitable adjustments and modifications to it.

It was left to Albert Einstein to fully appreciate the real importance of the failure of all ether-drift experiments. From the negative results of the ether-drift experiments and from his own reasoning. Einstein felt fully convinced that motion through ether is a meaningless concept, only motion relative to material bodies alone has physical significance. He analysed the physical consequences implied in the absence of such an absolute or fixed frame of reference and had the boldness to break away from old and traditional concepts of space and time. He gave new interpretations of time and space which completely revolutionized old ideas and thus laid the foundation for his new theory of relativity which was able not only to overcome the difficulties met with previously in coordinating experimental facts but also to throw new light on all physical phenomena.

There are two parts in the theory of relativity: (i) Special or restricted relativity and (ii) General relativity. The special theory of relativity, developed in 1905, treats problems involving inertial frames of reference, which are frames of reference moving at constant velocity with respect to one another. The general theory of relativity, the second and the more complex and difficult part of the theory, proposed by Einstein in 1915, treats problems involving frames of reference accelerated with respect to one another.

We shall concern ourselves here with the special theory of relativity which is not only comparatively simple but has also produced the most profound effect on the entire field of physics.

Special theory of relativity

The special theory of relativity is based upon two postulates.

1. *The laws of physics may be expressed in equations having the same form in all inertial frames of reference i.e., frames of reference moving at constant velocity with respect to one another.*

As can be readily seen, the first postulate expresses the absence of an absolute or fixed frame of reference. For, if the laws of physics were to take on different forms in different frames of reference, it could be determined from these differences as to which one of them are at rest in space and which are in motion. But because there is no absolute or fixed frame of reference, this distinction between state of rest and of uniform motion is not possible.

2. *The velocity of light in free space (vacuum) is a constant (same for all observers), independent not only of the direction of propagation but also of the relative velocity of the source and the observer.*

The second postulate follows directly from the result of the Michelson-Morley and other experiments.

The first postulate is a principle of equivalence which is generalized from a wide range of physical experience. It appears as a natural extension to all physical laws of the Newtonian principle of relativity which applied directly to mechanical laws. All physical processes are based on the fundamental equations of mechanics. Hence if the latter are the same with regard to all frames of reference that are in uniform motion relative to one another, then the equations of all physical phenomena must follow the same principle of equivalence. Thus the first postulate was an extension of the Newtonian principle of relativity to include not only the laws of mechanics but also those of rest of physics, including optics, electricity and magnetism.

The second postulate seems to be the hardest to accept, for it violates common sense notion. (Common sense, according to Einstein, is that layer of prejudices laid down in the mind prior to the age of eighteen). First of all, we have to think of light traveling in empty space. Giving up the ether is not too hard - after all it could never be detected. But the second postulate also tells us that the speed of light in vacuum is always the same - 3 x 10^8 metre per

second, no matter what the speed of the observer or the source. Thus a person moving toward or away from a source of light will measure the same speed for that light as someone at rest with respect to the source. This conflicts with our everyday notions, for we would expect to have to add in the velocity of the observer. Part of the problem is that in our everyday experience we do not measure velocities anywhere near the speed of light, thus we cannot expect our everyday experience to be helpful with the velocity of light. On the other-hand Michelson-Morley and other experiments are fully consistent with the second postulate.

By doing away with the idea of an absolute (universal) frame of reference, it was possible to reconcile Maxwell's electromagnetic theory with mechanics. The speed of light predicted by Maxwell's equations is the speed of light in vacuum in any frame of reference. The interrelationship of space and time coordinates also follows from the postulates of special theory of relativity which requires the giving up the common notions of space and time. Einstein's treatment presents a more advanced treatment of space and time than does the classical theory. The interdependence of space and time coordinates means that time must be treated in the same way as space coordinates. This gave rise to the idea of a four-dimensional space time continuum as well as to popular misconceptions about the *fourth dimension.*

Example 11.1 *In the Michelson-Morley experiment, the wavelength of monochromatic light used is 5000 Angstrom Units. What will be the expected fringe-shift on the basis of stationary ether hypothesis if the effective length of each path be 5 metres? (velocity of the earth = 3×10^4 m/sec, c = 3×10^8 m/sec and 1 A.U = 10^{-10}m).*

Soln.

Let the fringe shift be n fringes. Then

$$n = \frac{2Dv^2}{c^2\lambda}$$

Have v = 3×10^4m/sec

c = 3×10^8m/sec

D= 5 m

$\lambda = 5000 \times 10^{-10}\text{m} = 5 \times 10^{-7}\text{m}$

$$= \frac{2\times5\times(3\times10^4)^2}{(3\times10^8)^2\times5\times10^{-7}}$$

$= 0.2$ or $1/5$.

The expected fringe shift is one-fifth of a fringe-width.

11.3 Equations of relativity

Galilean - Newtonian transformations

Let us first consider the pre-Einstein relativity of physical quantities in classical or Galilean-Newtonian physics, and ask how events in one system S appear from another system S' the two systems being in uniform motion relative to one another (Fig. 11.2). These frames are referred to as *inertial frames* in which the law of inertia - Newton's first law, as well as second law of motion hold good. Now an event is something that happens independently of the reference frame that may be used to describe it. For concreteness, an

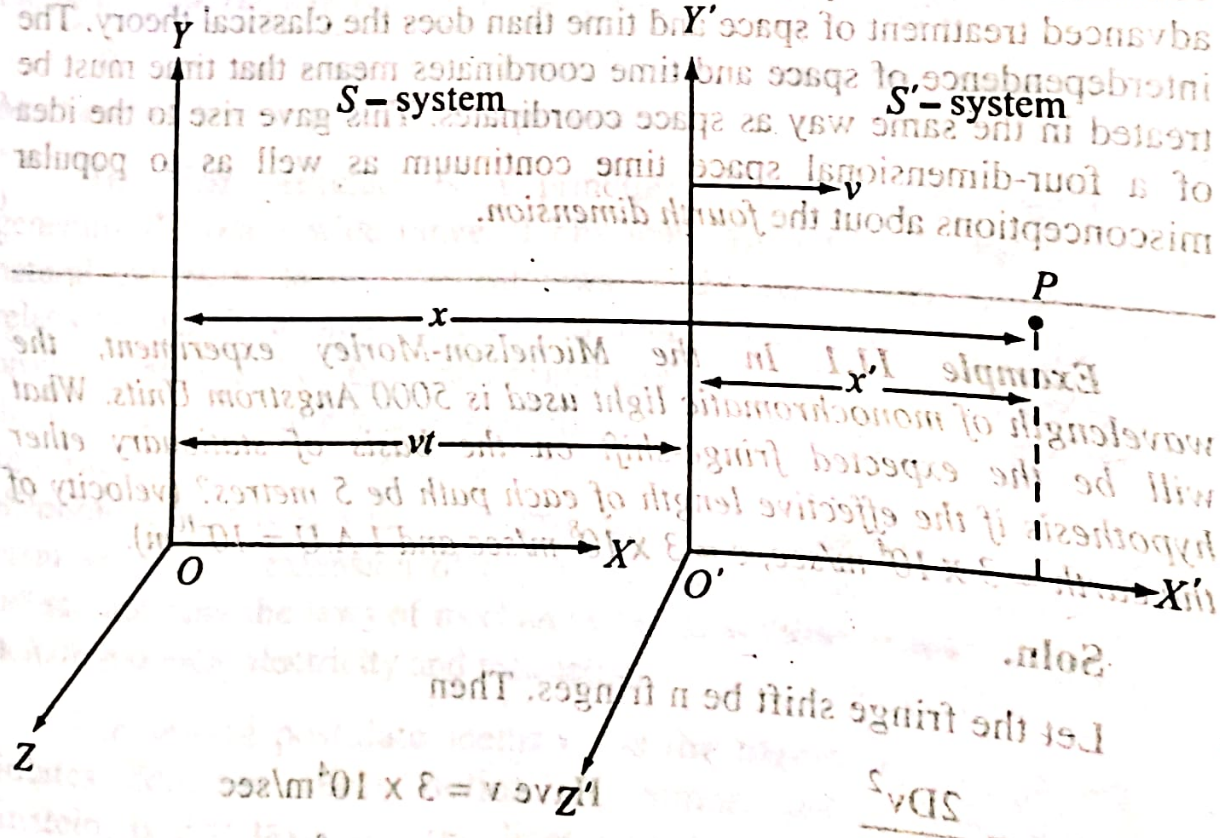


Fig. 11.2

event may be imagined to be a collision of two particles or the turning-on of a tiny light source. Let us suppose that we are in a frame of reference S. We specify an event by four (space-time) measurements in a particular frame of reference, say the position numbers x, y, z and the time t. An observer at another frame of reference S' which is moving (say in the +ve x direction) with a constant velocity v relative to S will find that the same event occurs at the time t' and has the coordinates x', y' and z'. How, are the measurements x, y, z, t related to x', y', z', t'?

At first glance the answer seems obvious enough. For simplicity, let us say that time in both systems is measured from the instant when the origins of S and S' coincided. Then measurements in x direction made in S will exceed those made in S' by an amount vt, which represents the distance that S' has moved in the x direction. Thus

$$x' = x - vt \tag{11.3a}$$

Since there is no relative motion in the y and z directions,

$$y' = y \tag{11.3b}$$

and

$$z' = z \tag{11.3c}$$

In classical physics it is assumed that time can be defined independently of any particular frame of reference. Therefore,

$$t' = t \tag{11.3d}$$

The set of equation (11.3a) to (11.3d) is known as Galilean or Newtonian transformation.

Let the positions of the end points of a rod lying in the S' frame of reference as recorded by an observer in S' be x_1' and x_2'. The same position as recorded by an observer in S are x_1 and x_2. Then the length of the same rod as recorded by the observer in S is

$$x_2 - x_1$$

The length of the same rod as recorded by the observer in S' is

$$x_2' - x_1' = (x_2 - vt) - (x_1 - vt)$$

$$= x_2 - x_1$$

the same as recorded by the observer in S.

Thus the length or *distance* is an *invariant quantity* when transformed from one coordinate system to another.

How do the measurements of different inertial observers compare with regards to velocities and accelerations of objects? Since the position of a particle in motion is a function of time, velocity and acceleration of a particle can be expressed in terms of time derivatives of position. We need only carry out successive time differentiations of the Galilean transformations. The velocity transformation follows at once.

Starting from x' = x – vt differentiation with respect to t gives

$$\frac{dx'}{dt} = \frac{dx}{dt} - v$$

But because t = t', the operation d/dt is identical to the operation d/dt', so that

$$\frac{dx'}{dt} = \frac{dx'}{dt'}$$

Therefore,

$$\frac{dx'}{dt'} = \frac{dx}{dt} - v$$

Similarly,

$$\frac{dy'}{dt'} = \frac{dy}{dt}$$

and

$$\frac{dz'}{dt'} = \frac{dz}{dt}$$

But $\frac{dx'}{dt'} = u'_x$, the x-component of the velocity measured in S', and dx/dt = u_x the x-component of the velocity measured in S. Therefore,

$$u'_x = u_x - v$$

$$u'_y = u_y$$

$$u'_z = u_z$$

In more general case in which v, the relative velocity of the frames, has components along all three axes, the more general result will be

$$u' = u - v \tag{11.4}$$

Thus the *velocities* measured by the two observers are not the same. In other words, they are *not invariant* under a transformation between Galilean-Newtonian coordinate systems.

The acceleration transformation can be obtained by merely differentiating the velocity relations. Proceeding as before, we obtain

$$\frac{d}{dt'}(u'_x) = \frac{d}{dt}(u_x - v)$$

or $\dfrac{du'_x}{dt'} = \dfrac{du_x}{dt}$ v being a constant

$$\frac{du'_y}{dt'} = \frac{du_y}{dt}$$

and $\dfrac{du'_z}{dt'} = \dfrac{du_z}{dt}$

That is, $a_x' = a_x$, $a_y' = a_y$ and $a_z' = a_z$. Hence $a' = a$. The measured components of acceleration of a particle are unaffected by the uniform relative velocity of the reference frames. In other words, acceleration are invariant under transformation between Galilean-Newtonian inertial frames of reference.

The Galilean-Newtonian transformation and the velocity transformation it leads to are both in accord with our intuitive expectations. However, they violate both the postulates of special theory of relativity. The first postulate calls for identical equations of physics in both S and S' frames of reference, but the fundamental equations of electricity and magnetism assume very different forms when eqns. 11.3a to 11.3d are used to convert quantities measured in one frame into their equivalents in the other. According to second postulate the speed of light c should have the same value whether

determined in S or S'. But if we measure the speed of light in the x direction in the system S to be c, then according to velocity transformation (eqn. 11.4), the speed c' in the S' system will be

$$c' = c - v$$

Clearly a different set of transformation is required if the postulates of the special theory of relativity are to be satisfied.

11.4 Lorentz transformation

Let us go back to our two observers at O and O' in systems S and S'. S' is moving with a constant velocity v relative to system S along the positive x direction. Suppose we make measurements of time from the instant when the origins of S and S' just coincide *i.e.*, t = 0 when O and O' coincide. Let a light pulse be emitted at t = t' = 0 *i.e.*, when O and O' coincide. The light pulse emitted will spread out as a spherical wavefront of ever increasing radius. Since at time t = 0, each observer was at the source of a spherical light wave, each observer must feel that he is at the centre of the growing sphere of light. Clearly, equation of the wavefront, as seen by the observer at O in the reference frame S, is

$$x^2 + y^2 + z^2 = c^2t^2 \qquad \text{(i)}$$

And equation of the wavefront, as seen by the observer at O' in reference frame S', is

$$x'^2 + y'^2 + z'^2 = c^2t'^2 \qquad \text{(ii)}$$

where c, the velocity of light is the same in either frame of reference.

In order that both the observers may seem to be at the centre of the same expanding wavefront relation (i) should be equal to relation (ii).

Now Galilean transformation gives x' = x – vt, y' = y, z' = z and t' = t. Substituting for x', y', z' and t' in relation (ii) above, we have

$$x^2 - 2xvt + v^2t^2 + y^2 + z^2 = c^2t^2$$

which is not the same as relation (i). Thus the Galilean transformation fails once again if the constancy of the value of c be assumed.

This clearly means that in the new transformation we are looking for, the relation x' = x – vt no longer holds good. Let a reasonable guess as to the kind of relationship between x and x' be

$$x' = k(x - vt) \qquad \text{(iii)}$$

where k is a factor of proportionality that does not depend upon either x or t but may be a function of v. The choice of eqn. (iii) follows from several considerations.

(a) It is simple and a simple solution to a problem should always be explored

(b) It is linear in x or x' *i.e.*, it will automatically give only one value of x' for a given value of x without having to impose any further conditions. This means a single event in frame S corresponds to a single event in frame S' as it must. A quadratic equation, on the other hand, will obviously give two interpretations, which is simply inadmissible.

(c) It has the possibility of reducing to eqn. $x' = x - vt$ which we know to be quite in accord with classical mechanics.

Because the equations of physics must have the same form in both S and S', we need only change the sign of v (in order to take into account the difference in direction of relative motion) to write the corresponding equation for x in terms of x' and t', or

$$x = k(x' + vt') \qquad \text{(iv)}$$

The factor k must be same in both frames of reference since there is no difference between S and S' other than in the sign of v.

As in the case of the Galilean transformation, there is nothing to indicate that there might be differences between the corresponding coordinates y, y' and z, z' which are normal to the direction of v. Hence we again write

$$y' = y \qquad \text{(v)}$$

and $$z' = z \qquad \text{(vi)}$$

The possibility that the time coordinates t and t' are not equal is now taken into account. The key to the problem is that the two observers have different concept of time. Two events judged to be simultaneous by one observer are not simultaneous for the other.

The transformation equation for time is obtained by substituting the value of x' given by eqn. (iii) into eqn. (iv); we then obtain

$$x = k^2(x - vt) + kvt'$$

$$\text{or,}\quad kvt' = k^2vt + x(1 - k^2)$$

$$\text{or,}\quad t' = kt + \frac{1-k^2}{kv}.x$$

$$\text{or,}\quad t' = k\left[\frac{x}{v}\left(\frac{1}{k^2}-1\right)+t\right] \qquad \text{(vii)}$$

Substituting eqns. (iii) and (v) into eqn. (ii) we obtain

$$k^2(x - vt)^2 + y^2 + z^2 = c^2k^2\left[\frac{x}{v}\left(\frac{1}{k^2}-1\right)+t\right]^2$$

Expanding and collecting terms we have

$$\left[k^2 - \frac{c^2k^2}{v^2}\left(\frac{1}{k^2}-1\right)\right]^2 x^2 - \left[2vk^2 + \frac{2c^2k^2}{v^2}\left(\frac{1}{k^2}-1\right)\right] xt$$

$$+ y^2 + z^2 = [c^2k^2 - v^2k^2]\, t^2 \qquad \text{(viii)}$$

Eqn. (vi) must be identical to eqn. (i). For this to be the case, the quantities in brackets in eqn. (vi) must be equal to 1,0, and c^2 respectively. It can easily be seen that all three requirements are fulfilled if

$$k = \frac{1}{\sqrt{1-v^2/c^2}}$$

For example, let us take the case of

$$2vk^2 + \frac{2c^2k^2}{v}\left(\frac{1}{k^2}-1\right) = 0$$

$$\text{or,}\quad \frac{2ck^2}{v}\left(\frac{1}{k^2}-1\right) = -2vk^2$$

$$\text{or,}\quad \frac{1}{k^2}-1 = -\frac{2vk^2}{2c^2k^2}\cdot v = -\frac{v^2}{c^2}$$

$$\text{or,} \quad \frac{1}{k^2} = 1 - \frac{v^2}{c^2}$$

$$\text{or,} \quad k = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \text{(ix)}$$

Insertion of the above value of k in eqns. (iii), (v), (vi) and (vii) lead us to the complete transformations of the measurement of one event made in S to the corresponding measurements in S'. Or,

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \qquad (11.5a)$$

$$y' = y \qquad (11.5b)$$

$$z' = z \qquad (11.5c)$$

$$t' = \frac{t - \frac{v}{c^2} . x}{\sqrt{1 - v^2/c^2}} \qquad (11.5d)$$

In obtaining these transformations, the positive root has been chosen, so that $v = 0$ implies $x' = x$ and $t' = t$.

These transformations are referred to as *Lorentz transformation*, after H.A. Lorentz, a Dutch physicist, because it was he who first showed that this transformation was the only one under which the laws of electricity and magnetism have the same form in all reference frames in relative motion (*i.e.*, in all inertial frames). It was not until a number of years latter that Einstein showed their full significance as being the only proper transformation for all types of measurements, electromagnetic or otherwise.

We have, in our discussion above, assumed S' to be moving in the positive x-direction with velocity v relative to S. As seen from S', therefore, S would appear to be moving in negative x-direction, *i.e.*, with velocity –v. Thus in order to transform measurements made in frame S' into those in frame S, we need only to replace v by -v. We shall then have

$$x = \frac{x' + vt'}{\sqrt{1 - v^2/c^2}} \quad (11.6a)$$

$$y = y' \quad (11.6b)$$

$$z = z' \quad (11.6c)$$

$$t = \frac{t' + \frac{v}{c^2}x'}{\sqrt{1 - v^2/c^2}} \quad (11.6d)$$

These transformations are known as Inverse Lorentz transformations.

As will be readily seen that when $v << c$ (*i.e.*, when $v/c \rightarrow 0$) and $x << ct$, the Lorentz (or relativistic) transformations reduce to Galilean or Newtonian transformations. Thus Galilean or Newtonian physics is just a particular case of relativistic physics. In fact, the tremendous velocity of earth round the sun is 18 miles per sec., but this is still only 1/10000[th] of the velocity of light. The Galilean or Newtonian physics is, therefore, perfectly valid for almost all our practical purposes, the relativistic effects manifesting themselves only at very high values of v, when they are comparable with c, as in the case of electrons, protons, mesons, etc.

Lorentz transformations, though they differ only slightly from Galilean transformations in most physical situations, bring in a profoundly new concept in the kinematics of the universe. They remove the universal concept of time and treat them on the same footing as space coordinates. The time in the second frame S' is not the same as that of S (except when v<<c): it depends not only on the relative velocity v but also on x - the abscissa in the first frame of reference. Hence, according to theory of relativity, *events which happen at the same place at different times as viewed from one reference frame may be seen from another frame to happen at different places as well.* Similarly a difference in *spatial* position with respect to one system may correspond to a difference in both space and time with respect to another. Thus a space difference can be converted partly into a time difference and vice versa by a mere change in the *frame of reference* that is used. For this reason space and time are considered as two intimately connected aspects of a *four-dimensional* space-time continuum.

Another aspect of Lorentz transformation is that, v must be less than c, otherwise $\sqrt{1-v^2/c^2}$ becomes imaginary. Hence in theory no body can move in vacuum with a velocity greater than that of light.

11.5 Relativistic velocity transformation

Consider a ray of light to be emitted in a frame of reference S' in the direction of its motion. The frame S' is moving with a velocity v relative to another frame of S which is at rest. Then the speed of light as measured in S should be c + v. But this is against the second postulate of the special theory of relativity according to which the speed of light has the same value for all observers whether at rest or in relative motion. The problem can be solved by considering Lorentz transformation equations.

Let us consider that something moves relative to both S and S'. The three components of its velocity as measured by an observer in the frame S are

$$V_x = \frac{dx}{dt}$$

$$V_y = \frac{dy}{dt}$$

$$V_z = \frac{dz}{dt}$$

while to an observer in S', these components are

$$V'_x = \frac{dx'}{dt'}$$

$$V'_y = \frac{dy'}{dt'}$$

$$V'_z = \frac{dz'}{dt'}$$

Differentiating the Lorentz transformation equations for x', y', z' and t', we obtain

$$dx' = \frac{dx - vdt}{\sqrt{1-v^2/c^2}}$$

$$dy' = dy$$

$$dz' = dz$$

$$dt' = \frac{dt - \frac{vdx}{c^2}}{\sqrt{1-v^2/c^2}}$$

Therefore,

$$V'_x = \frac{dx'}{dt'}$$

$$= \frac{dx - vdt}{dt - \frac{vdx}{c^2}}$$

$$= \frac{\frac{dx}{dt} - v}{1 - \frac{v}{c^2}\frac{dx}{dt}}$$

$$= \frac{V_x - v}{1 - \frac{vV_x}{c^2}} \qquad (11.7a)$$

$$V'_Y = \frac{dy'}{dt'}$$

$$= \frac{dy\sqrt{1-v^2/c^2}}{dt - \frac{vdx}{c^2}}$$

$$= \frac{\frac{dy}{dt}\sqrt{1-v^2/c^2}}{1 - \frac{v}{c^2}\frac{dx}{dt}}$$

$$= \frac{V_y\sqrt{1-v^2/c^2}}{1-\frac{vV_x}{c^2}} \qquad (11.7b)$$

Similarly,

$$V'_Z = \frac{V_z\sqrt{1-v^2/c^2}}{1-\frac{vV_x}{c^2}} \qquad (11.7c)$$

Eqns. 11.7a to 11.7c constitute the relativistic velocity transformation equations. When v is small compared to c, these equations reduce to the classical equations.

$$V'_X = V_X - v$$

$$V'_Y = V_Y$$

$$V'_Z = V_Z$$

V_X, V_Y, V_Z, can similarly be expressed in terms of V'_X, V'_Y and V'_Z by replacing v with -v and exchanging primed and unprimed quantities in eqns. 11.7a to 11.7c. Thus the inverse velocity transformation equations are

$$V_X = \frac{V_{x'} + v}{1+\frac{vV'_x}{c^2}} \qquad (11.8a)$$

$$V_Y = \frac{V'_y\sqrt{1-v^2/c^2}}{1+\frac{vV'_x}{c^2}} \qquad (11.8b)$$

$$V_Z = \frac{V'_z\sqrt{1-v^2/c^2}}{1+\frac{vV'_x}{c^2}} \qquad (11.8c)$$

Let a ray of light be emitted in the moving frame S' in the direction of its motion relative to the stationary frame S. Then the velocity of light as measured by an observer in the frame S is given by

$$V_x = \frac{V'_x + v}{1 + \frac{vV'_x}{c^2}}$$

In this case $V'_x = c$

$$\therefore \quad V_x = \frac{c + v}{1 + \frac{c.v}{c^2}} = \frac{c(c+v)}{c+v} = c.$$

Thus both observers measure the same value for the speed of light, as they must according to the second postulate of the special theory of relativity.

Example 11.2 Show by means of Lorentz transformation equations that

$$x'^2 - c^2t'^2 = x^2 - c^2t^2$$

Soln.

In accordance with Lorentz transformation we have

$$x' = k(x - vt) \text{ and } t' = k\left(t - \frac{vx}{c^2}\right)$$

$$\text{where } k = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$\therefore x'^2 - c^2t'^2 = k^2(x - vt)^2 - c^2k^2\left(t - \frac{vx}{c^2}\right)^2$$

$$= k^2(x^2 - 2xvt + v^2t^2) - c^2k^2\left(t^2 - \frac{2vxt}{c^2} + \frac{v^2x^2}{c^4}\right)$$

$$= k^2x^2\left(1 - \frac{v^2}{c^2}\right) - k^2c^2t^2\left(1 - \frac{v^2}{c^2}\right)$$

$$= k^2\left(1-\frac{v^2}{c^2}\right)\left(x-c^2t^2\right)$$

$$= \frac{1-\dfrac{v^2}{c^2}}{1-\dfrac{v^2}{c^2}}\left(x-c^2t^2\right) = x^2 - c^2t^2.$$

***Example 11.3** Show that if (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) are the coordinates of one event in S_1 and the corresponding event in S_2 respectively, then the expression*

$$ds_1^2 = dx_1^2 + dy_1^2 + dz_1^2 - c^2dt_1^2$$

is invariant under a Lorentz transformation of coordinates (i.e., $ds_1^2 = ds_2^2$).

Soln.

The Inverse Lorentz transformation equations are given by

$$x_1 = \frac{x_2 + vt^2}{\sqrt{1-\beta^2}};\ y_1 = y_2;\ z_1 = z_2;\ \text{and } t_1 = \frac{t_2 + \dfrac{vx_2}{c^2}}{\sqrt{1-\beta^2}}$$

where $\beta = v/c$

Differentiating,

$$dx_1 = \frac{dx_2 + vdt_2}{\sqrt{1-\beta^2}};\ dy_1 = dy_2;\ dz_1 = dz_2;\ \text{and } dt_1 = \frac{dt_2 + \dfrac{v}{c^2}.dx_2}{\sqrt{1-\beta^2}}$$

$$= \frac{dt_2 + (\beta/c)dx_2}{\sqrt{1-\beta^2}}$$

$$\therefore\ ds_1^2 = dx_1^2 + dy_1^2 + dz_1^2 - c^2dt_1^2$$

$$= \left(\frac{dx_2 + vdt_2}{\sqrt{1-\beta^2}}\right)^2 + dy_2^2 + dz_2^2 - c^2\left[\frac{dt_2 + (\beta/c)dx_2}{\sqrt{1-\beta^2}}\right]^2$$

$$= \frac{1}{(1-\beta^2)}(dx_2 + vdt_2)^2 - \frac{c^2}{(1-\beta^2)}[dt_2 + (\beta/c)dx_2]^2 + dy_2^2 + dz_2^2$$

$$= \frac{1}{(1-\beta^2)}\Big[dx_2^2 + 2vdx_2dt_2 + v^2dt_2^2 - c^2dt_2^2 - c^2.2.\frac{\beta}{c}dx_2dt_2$$

$$-\beta^2dx_2^2\Big] + dy_2^2 + dz_2^2$$

$$= \frac{1}{(1-\beta^2)}\left[dx_2^2 + v^2dt_2^2 - c^2dt_2^2 - \beta^2dx_2^2\right] + dy_2^2 + dz_2^2$$

$$= \frac{1}{(1-\beta^2)}\left[dx_2^2(1-\beta^2) + dt_2^2(v^2 - c^2)\right] + dy_2^2 + dz_2^2$$

$$= dx_2^2 + dt_2^2\frac{(v^2 - c^2)}{1 - \frac{v^2}{c^2}} + dy_2^2 + dz_2^2$$

$$= dx_2^2 - c^2.dt_2^2\frac{(v^2 - c^2)}{(v^2 - c^2)} + dy_2^2 + dz_2^2$$

$$= dx_2^2 + dy_2^2 + dz_2^2 - c^2dt_2^2$$

$$= ds_2^2.$$

Example 11.4 *What would be the speed of a passenger who wishes to pass a spaceship, whose speed with respect to earth is 0.9c, at a relative speed of 0.5c?*

Soln.

According to conventional mechanics, the actual speed of the passenger is 0.9c + 0.5c = 1.4c – more than the velocity of light.

But according to relativistic mechanics, the required speed of the passenger is

$$U = \frac{u' + v}{1 + \frac{u'v}{c^2}}$$

Here $v = 0.9c$

$u' = 0.5c$

$$= \frac{0.9c + 0.5c}{1 + \frac{(0.9c)(0.5c)}{c^2}}$$

$$= 0.9655c.$$

The passenger must move with a speed of 0.9655c in order to have a relative speed of 0.5c with respect to the spaceship.

***Example 11.5** Two electrons leave a radioactive sample in opposite directions, each having a speed 0.67c with respect to the sample. What is the speed of one electron relative to the other according to (i) ordinary mechanics and (ii) relativistic mechanics?*

Soln.

In ordinary mechanics, the relative speed is 0.67c + 0.67c = 1.34c - greater than the velocity of light.

According to relativistic mechanics,

$$u = \frac{u' - v}{1 - \frac{u'v}{c^2}}$$

Here $v = -0.67c$

$u' = 0.67c$

$$= \frac{(0.67 + 067)c}{1 + (0.67)}$$

$$= \frac{1.34}{1.45}c = 0.92c.$$

Thus the speed of one electron relative to the other is less than c.

***Example 11.6** A photon is traveling east and another photon is traveling west. Find the relative velocity of the two photons.*

$$u = \frac{u' - v}{1 - \frac{u'v}{c^2}}$$

Here u' = c, v = − c

$$\therefore \quad u = \frac{c + c}{1 + \frac{c \times c}{c^2}} = c.$$

***Example 11.7** The interval S_{12} between two events is defined by the relation $S_{12}^2 = x^2 + y^2 + z^2 + w^2$, where $w = ict$ and $i = \sqrt{-1}$. Show with the help of Lorentz transformation that S_{12} has the same value in all inertial frames, though distance and time may have different values.*

Soln.

By Lorentz transformation, we have

$$S_{12}'^2 = x'^2 + y'^2 + z'^2 - c^2t'^2$$

$$= \frac{(x - vt)^2}{1 - \frac{v^2}{c^2}} + y^2 + z^2 - c^2 \frac{\left(t - \frac{vx}{c^2}\right)^2}{1 - \frac{v^2}{c^2}}$$

$$= k^2(x - vt)^2 + y^2 + z^2 - k^2c^2 \left(t - \frac{vx}{c^2}\right)^2 \left[k = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\right]$$

$$= k^2\left[x^2 - 2xvt + v^2t^2\right] + y^2 + z^2 - c^2k^2\left[t^2 - \frac{2xvt}{c^2} + \frac{v^2x^2}{c^4}\right]$$

$$= k^2\left[x^2 - 2xvt + v^2t^2\right] + y^2 + z^2 - k^2\left[c^2t^2 - 2xvt + \frac{v^2x^2}{c^2}\right]$$

$$= k^2\left[x^2 - 2xvt + v^2t^2 - c^2t^2 + 2xvt - \frac{v^2x^2}{c^2}\right] + y^2 + z^2$$

$$= k^2\left[x^2 - \frac{v^2x^2}{c^2} - c^2t^2 + v^2t^2\right] + y^2 + z^2$$

$$= \frac{c^2}{c^2 - v^2}\left[x^2 - \frac{v^2x^2}{c^2} + c^2t^2 + v^2t^2\right] + y^2 + z^2$$

$$= \frac{1}{c^2 - v^2}\left[c^2x^2 - v^2x^2 - c^4t^2 + c^2v^2t^2\right] + y^2 + z^2$$

$$= \frac{1}{c^2 - v^2}\left[x^2(c^2 - v^2) - c^2t^2(c^2 - v^2)\right] + y^2 + z^2$$

$$= \frac{1}{c^2 - v^2}\left[(c^2 - v^2)(x^2 - c^2t^2)\right] + y^2 + z^2$$

$$= x^2 - c^2t^2 + y^2 + z^2$$

$$= x^2 + y^2 + z^2 - c^2t^2$$

$$= x^2 + y^2 + z^2 + (ict)^2$$

$$= x^2 + y^2 + z^2 + w^2$$

$$= S_{12}^2$$

or, $S'_{12} = S_{12}$

Thus S_{12} has same value in all inertial frames, though distance and time may have different values.

Example 11.8 Obtain the resultant of two velocities of 0.6 c and 0.8 c respectively, inclined to one another at an angle of 60°.

Soln.

Let one of the velocities, say 0.6 c, be imagined to be inclined along the x-axis while the other, 0.8 c, be inclined to it at an angle of 60°. This is equivalent to saying that a frame of reference S' is moving along the x-axis with a velocity 0.6 c relative and parallel to

stationary frame of reference S and a particle is moving with a velocity of 0.8 c inclined to it at an angle of 60°. It is then required to determine the resultant velocity of the particle as it would appear to an observer in frame S.

Component of velocity 0.8c along x-axis,

$$U'_x = 0.8c \cos 60° = 0.8c \times \frac{1}{2} = 0.4c.$$

Component of 0.8c along y-axis,

$$U'_y = 0.8c \sin 60° = 0.8c \times \frac{\sqrt{3}}{2} = 0.4\sqrt{3}.c.$$

Applying Lorentz transformation, the component along x-axis.

$$U_x = \frac{U'_x + v}{1+\frac{U'_x.v}{c^2}} = \frac{0.4c+0.6c}{1+\frac{0.4c \times 0.6c}{c^2}} = 0.8064\ c.$$

The component along y-axis.

$$U_y = \frac{U'_y}{k\left(1+U'_x.\frac{v}{c^2}\right)} \qquad \text{where } k = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{or, } U_y = \frac{U'_y\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{U'_x.v}{c^2}}$$

$$= \frac{0.4\sqrt{3}.c\sqrt{1-\frac{(0.6c)^2}{c^2}}}{1+0.4c\frac{0.6c}{c^2}}$$

$$= 0.447\ c.$$

The resultant velocity U of the particle as observed in frame S

$$= \sqrt{U_x^2 + U_y^2}$$

$$= \sqrt{(0.8064c)^2 + (0.447c)^2}$$

$$= 0.922c$$

$$= 0.922 \times 3 \times 10^8$$

$$= 2.766 \times 10^8 \text{ m/sec.}$$

The angle of inclination,

$$\theta = \tan^{-1}\frac{U_y}{U_x}$$

$$= \tan^{-1}\frac{0.477}{0.8064}$$

$$= 29°15' \text{ with velocity } 0.6c.$$

11.6 Some consequences of the Lorentz transformation equations - length contraction and time dilation.

The Lorentz transformation equations derived in Art. 11.4, have some interesting consequences for length and time measurements. These will appear to be surprisingly new and strange because, in view of the small relative motion between the frames of reference of which we have experience in our daily life, we do not ordinarily come across any perceptible relativistic phenomena.

Let us consider a few important cases:

(i) *Length contraction*: From Lorentz transformation equations it can be shown that length is not absolute but relative. Let a rod be lying (at rest) along the x'-axis of a moving frame of reference S'. An observer in this frame of reference S' determines the co-ordinates of the ends of the rod to be x'_1 and x'_2 and concludes that the length l_0 of the rod is

$$l_0 = x'_2 - x'_1$$

l_0 is the length of the rod in a moving frame of reference S' but the rod itself is at rest. This length of the rod as measured by an

observer at rest with respect to the rod is called its *proper length*. What is the length of the rod measured by an observer in the frame of reference S relative to which the rod (or the frame) is moving with a velocity v?

Let the observer in the S frame determine the coordinates of the ends of the same rod lying at rest in the moving frame S' to be x_1 and x_2. Then the length l of the rod as measured by an observer in S is

$$l = x_2 - x_1$$

Now applying Lorentz transformation we have

$$x'_2 = \frac{x_2 - vt}{\sqrt{1 - v^2/c^2}}$$

and

$$x'_1 = \frac{x_1 - vt}{\sqrt{1 - v^2/c^2}}$$

and so

$$l_0 = x'_2 - x'_1$$

$$= \frac{x_2 - x_1}{\sqrt{1 - v^2/c^2}}$$

$$= \frac{l}{\sqrt{1 - v^2/c^2}}$$

or, $$l = l_0\sqrt{1 - v^2/c^2} \qquad (11.9)$$

Since $\sqrt{1 - v^2/c^2}$ is always less than unity, l will always be less than the proper length l_0, and therefore it is said that the length has contracted. Thus the Fitzgerald - Lorentz contraction discussed in Art. 11.1 is mathematically the same as the relativistic length contraction given by eqn. 11.9, although the concepts underlying the two contraction are different.

Thus the length of a body is measured to be greatest when it is at rest relative to the observer. When it moves with a velocity v

relative to the observer its measured length is contracted by the factor $\sqrt{1-v^2/c^2}$. Stated in more simple terms: *the length of an object is measured to be shorter when it is moving than when it is at rest.*

It should be emphasized that because the relative velocity of the two frames S and S' appears only as v^2 in $\sqrt{1-v^2/c^2}$, it does not matter which frame we call S and which S'. An observer in frame S' will find that a rod at rest in frame S, parallel to the x-axis, is shortened by the same factor.

It is important to note, however, that the length contraction occurs only in the direction of motion whereas its dimensions perpendicular to the direction of motion are unaffected. Thus if we have two frames of reference in relative motion (one stationary and the other in motion) along the x-direction, say, a straight line parallel to this direction in one appears shorter to an observer in the other, - not so, however, a straight line along the y or the z-direction, perpendicular to the x-direction. Similarly, a square and a circle in one appear to an observer in the other to be a rectangle and an ellipse respectively, the sides of the square and the radius of the circle (or their components) in the direction of motion getting shortened (Fig. 11.3).

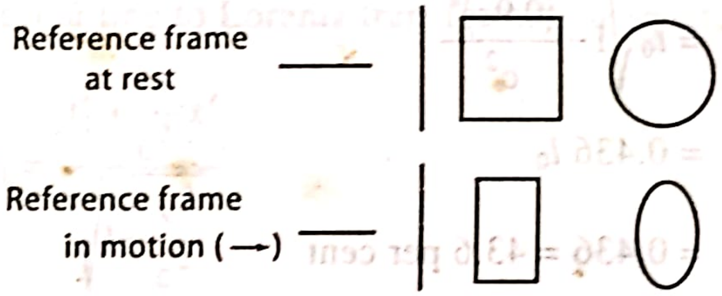


Fig. 11.3

It may be mentioned that, for small values of v for which $v/c \to 0$, the factor $\sqrt{1-v^2/c^2} = 1$, so that $l = l_0$. This is in accordance with classical mechanics where length is treated as an absolute quantity, unaffected by rest or motion. This is the reason why Fitzerald-Lorentz contraction is nil or negligible at speeds with which we are ordinarily concerned. The contraction becomes appreciable only when the velocity v is comparable to the velocity of light. For

example, let us suppose that a body is moving with a speed of 1000 miles/sec, an enormous velocity according to our everyday experience. Then the length as measured by an observer at rest is

$$l = l_0 \sqrt{1 - \frac{v^2}{c^2}}$$

$$= l_0 \sqrt{1 - \frac{(1{,}000)^2}{(1{,}86{,}000)^2}}$$

$$= 0.999985\, l_0$$

or, $\frac{l_0}{l}$ = 0.999985 = 99.9985 per cent

The contraction is, therefore, negligibly small and hence cannot be appreciated even when the body moves with a velocity of 1000 miles/sec.

Now suppose v = 0.9 c.

Then,

$$l = l_0 \sqrt{1 - \frac{(0.9c)^2}{c^2}}$$

$$= 0.436\, l_0$$

or, $\frac{l_0}{l}$ = 0.436 = 43.6 per cent

The contraction is, therefore, 43.6 per cent - a significant change.

(ii) *Time dilation* : The most remarkable consequence of Lorentz transformation equations is that time intervals too, like length, are affected by relative motion. A clock appears to an observer to run slower when it is in motion, then when it is at rest with respect to the observer. This means that the time interval between two events occurring at a given point in space in a frame of reference S', moving with respect to the observer in the frame S, appears greater as noted on a clock at rest with respect to the observer (frame S) than on an *identical clock* in the moving frame of

reference itself (S'). This *apparent lengthening of time* is referred to as time dilation. Let us deduce this result from Lorentz transformation.

Let the frame of reference S' be moving with a velocity v relative to the stationary frame of reference S along the positive x-direction. Let us further suppose that the two reference frames carry two identical clocks at their origin O and O' and that both the clocks show zero time when they just cross each other *i.e.*, the two origins just coincide.

Imagine a gun placed at the point x'_1 in the frame S'. Suppose it fires two shots at times t'_1 and t'_2 as noted on the clock carried by the frame S'. Let t_1 and t_2 be the corresponding times as noted on the clock carried by the frame S. Then the time interval between the two events (firing of the shots), as noted on the clock in the moving frame S' is given by

$$\Delta t' = (t'_2 - t'_1)$$

And the time interval between the same two events as noted on the clock in the stationary frame S is given by

$$\Delta t = (t_2 - t_1)$$

Now according to Lorentz transformation equations we have

$$t_1 = \frac{t'_1 + \frac{v}{c^2}x'}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{and } t_2 = \frac{t'_2 + \frac{v}{c^2}x'}{\sqrt{1-\frac{v^2}{c^2}}}$$

Therefore,

$$\Delta t = t_2 - t_1$$

$$= \frac{t'_2 - t'_1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$= \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (11.10)$$

Since $1/\sqrt{1-v^2/c^2}$ is greater than unity, we have $\Delta t > \Delta t'$.

Thus the time interval $\Delta t'$ between two events occurring at a given point in the moving frame S' appears to be longer, or dilated by a factor $1/\sqrt{1-v^2/c^2}$ to the observer in the stationary frame S.

The same will happen if the frame S' is at rest and S is in motion. Thus, time dilation like length contraction shows reciprocity effect *i.e.*, is independent of the direction of velocity and depends only on its magnitude. Stated in general terms, a clock is measured to go at its fastest rate when it is at rest relative to the observer. When it moves with a velocity v relative to the observer, its rate is measured to have slowed down by a factor $1/\sqrt{1-v^2/c^2}$.

It is not that the clocks are somewhat at fault. To the contrary, the clocks are assumed to be good ones. Time is actually measured to pass more slowly in any moving reference frame as compared to a reference frame at rest.

The time interval $\Delta t'$ between two events occurring at the same place, measured on the clock moving with the reference frame S' in which the events occur, and hence at rest with respect to it, is called the *proper time interval* or *proper time* and is usually denoted by τ. It is always less than the corresponding time interval, measured on a system at rest. The relation $\Delta t = \Delta t'/\sqrt{1-v^2/c^2}$ is true only when $\Delta t'$ represents the time interval between two events in a reference frame where the two events occur at the same points in space.

Time dilation has aroused interesting speculation about space travel. Under the old concept of time it would not be possible for

ordinary mortals to reach a star 100 light years away. One light year is the distance light can travel in one year = 3.0×10^8 m/sec x 3.15×10^7 sec = 9.5×10^{15}m. Even if a space ship could travel with the velocity of light, it would take 100 years to reach such a star. But time dilation tells us that the time involved would be less for an astronaut. In a space ship traveling at v = 0.999c, the time for such a trip would be only $t' = t\sqrt{1-v^2/c^2} = 100 \text{ yrs.} \sqrt{1-(0.999)^2} = 4.5 \text{ yrs.}$

Thus a person could make such a trip. Time dilation allows such a trip, but the enormous practical problems of achieving such speeds will not be overcome in foreseeable future.

It may be noticed that in the above example while 100 years would pass on earth, only 4.5 yrs would pass for the astronaut on the trip. Is it just the clocks that would slow down for the astronaut? The answer is no. Along with time all physiological processes including the process of ageing itself, are slowed down in a fast moving reference frame. While people on earth would experience 100 years of ordinary activity, the astronaut would experience 4.5 years of normal sleeping, eating, reading and so on.

Time dilation has given rise to a paradox called the *twin paradox.* According to this paradox one of a pair of 20 years old twins takes off in a spaceship traveling at a very high speed to a distant star and back, while the other twin remains on earth. According to time dilation whereas 20 years might pass for the earth twin, perhaps only 1 years - depending on the speed of the spaceship, would pass for the traveling twin. Thus when the traveler returns, the earth bound twin could expect to be 40 years old whereas his twin would be only 21. This is the viewpoint of the twin on the earth. But what is the viewpoint of the traveling twin? Since *everything is relative*, all inertial frames are equally good. Can't the astronaut twin claim that since the earth is moving away at high speed, time would pass more slowly on earth and the earth twin would age less? This is the opposite of what the earth twin predicts. They cannot both be right, for after all the spaceship returns to earth and a direct comparison of ages and clocks can be made.

This is, however, not a paradox at all. The consequences of special theory of relativity - in this case, time dilation, can be applied only by observers in inertial reference frame. The earth is

such a frame - or nearly so. But the spaceship cannot be regarded as an inertial frame as it accelerates at the start and end of its trip and, more importantly, when it turns around at the far point of its journey. During these periods of acceleration, the spaceships predictions based on relativity are not valid. The twin on earth is on an inertial frame and can make valid predictions. Thus there is no paradox. The prediction of the earth twin that the traveling twin returns having aged less is the proper one. This was subsequently confirmed by Einstein's general theory of relativity which deals with accelerating reference frames.

Illustration of time dilation and length contraction – Meson decay.

A striking illustration of both the time dilation and length contraction occurs in the case of the decay of unstable particles called μ-mesons. These elementary (subatomic) particles, having a mass 215 times that of an electron, are produced in the upper reaches of the atmosphere as a result of collision between fast cosmic ray particles, arriving at the earth from space, and the air molecules. Although they do not all have the same velocity, the faster ones among them have a speed of 2.994×10^8 ms^{-1}, which is 0.998 c. A μ-meson decays into an electron an average of 2.2×10^{-6} sec after it comes into being *i.e.*, μ-meson's mean lifetime is 2×10^{-6} sec. Even with the high velocity of 0.998 c, the μ-meson can travel in 2.2×10^{-6} sec (mean life time of the meson) a distance of-only

$$y = vt_0$$

$$= 2.994 \times 10^8 \times 2.2 \times 10^{-6}$$

$$= 660 \text{ m}$$

This is the distance they can travel before they disintegrate into electron and two neutrinos each. And, yet they are found copiously in the laboratory 10km below the spot where they are produced. How is this possible?

This meson paradox can be resolved by using the results of the special theory of relativity. Let us examine the problem from the meson's frame of reference in which its mean lifetime is 2.2×10^{-6}, sec. While the meson lifetime is unaffected by its motion, its distance to the ground appears shortened by the factor

$$\frac{y}{y_0} = \sqrt{1 - \frac{v^2}{c^2}}$$

which means that while we, on the ground, measure the altitude at which the meson is produced as y_0, the meson sees it as y. Therefore, if y = 660 m, the maximum distance that the meson can travel *in its own frame of reference* at the speed of 0.998 c before decaying, then the corresponding distance y_0, as measured from our reference frame is

$$y_0 = \frac{y}{\sqrt{1 - \frac{v^2}{c^2}}} \text{ m}$$

$$= \frac{600}{\sqrt{1 - \frac{(0.998c)^2}{c^2}}} \text{ m}$$

$$= \frac{600}{\sqrt{1 - 0.996}} \text{ m}$$

$$= \frac{600}{0.063} \text{ m}$$

$$= 10479 \text{ m.}$$

Hence, despite their brief life spans, it is possible for the mesons to reach the ground from the considerable altitudes at which they are actually produced.

Now let the same problem be examined from the reference frame of the observer. The altitude at which the meson is produced, as measured from the ground, is y_0. But the lifetime of the meson as measured from the ground (observer's reference frame) has been *extended or dilated*, owing to the relative motion, to the value t.

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

$$= \frac{2.2 \times 10^{-6}}{\sqrt{1 - \frac{(0.998c)^2}{c^2}}}$$

$$= \frac{2.2 \times 10^{-6}}{0.063} \text{ sec}$$

$$= 34.92 \times 10^{-6} \text{ sec.}$$

Therefore, in 34.92×10^{-6} sec, a meson whose speed is 0.998 c, can travel a distance

$$y_0 = vt$$

$$= 2.994 \times 10^8 \times 34.92 \times 10^{-6} \text{m}$$

$$= 10455 \text{ m.}$$

the same distance as obtained before. Thus the two points of view give identical results.

Example 11.9 (a) S' (moving frame) observes that two events occur at the same place but are separated in time. S (stationary frame) will then declare that the two events occur in different places.

(b) S' observes that the two events occur at the same time but are separated in space. S will then declare that two events occur at different times.

In (a) assume the time separation in S' to be 10 minutes. What is the distance separation observed by S?

In (b) assume the distance separation in S' to be 25 metres; what is the time separation observed by S?

Soln.

Take v = 20 m/s and c = 3×10^8 m/s.

(a) Here $t'_2 - t'_1 = 10$ minutes = 600 seconds

$$x_2 - x_1 = ?$$

$$x_2 = \frac{x'_2 + vt'_2}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{and} \quad x_1 = \frac{x'_1 + vt'_1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\therefore \quad x_2 - x_1 = \frac{(x'_2 - x'_1) + v(t'_2 - t'_1)}{\sqrt{1-\frac{v^2}{c^2}}}$$

Here $x'_2 = x'_1$

$$\therefore \quad x_2 - x_1 = \frac{v(t'_2 - t'_1)}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \frac{20\times 600}{\sqrt{1-\left(\frac{20}{3\times 10^8}\right)^2}}$$

$$\simeq 12000 \text{ m}$$

$$\simeq 12 \text{ km.}$$

(b) Here $x'_2 - x'_1 = 25$ m, $t_2 - t_1 = ?$

$$t_2 = \frac{t'_2 + \frac{vx'_2}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \qquad t_1 = \frac{t'_1 + \frac{vx'_1}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\therefore \quad t_2 - t_1 = \frac{(t'_2 - t'_1) + [v(x'_2 - x'_1)]/c^2}{\sqrt{1-\frac{v^2}{c^2}}}$$

But $t'_2 = t'_1$

$$\therefore \quad t_2 - t_1 = \frac{v(x'_2 - x'_1)}{c^2\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \frac{20\times 25}{(3\times 10^8)^2\sqrt{1-\left(\frac{20}{3\times 10^8}\right)^2}}$$

$$\simeq 5.55 \times 10^{-15} \text{ sec.}$$

Example 11.10 Pions (charged particles of mass between that of the electron and the proton and of positive or negative electronic charge) are radioactive and, when brought to rest, their half-life is measured to be 1.77 x 10^{-8} secs. That is, half of the number present at any time have decayed 1.77 x 10^{-8} sec. later. A collimated pion beam moving at a speed of 0.99c is found to drop to half its original intensity 39m from the target. Does time dilation account for the measurements?

Soln.

(i) If the half-life is 1.77 x 10^{-8} sec and the speed is 0.99 x 3 x 10^{-8} = 2.97 x 10^{8} m/sec, the distance traveled over which half the pions in the beam should decay is

$$d = vt = 2.97 \times 10^{8}\ \text{m/sec} \times 1.77 \times 10^{-8}\ \text{sec} = 5.3\ \text{m}.$$

This contradicts the direct measurement of 39 m.

(ii) In the absence of any relativistic effect, the half-life would be measured to be the same for pions at rest and pions in motion - an assumption made in (i). In relativity, however, the non proper and proper half-lives are related by

$$t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

Here $t_0 = 1.77 \times 10^{-8}$ sec.

$v = 0.99\ c$

$$t = \frac{1.77\times10^{-8}}{\sqrt{1-\left(\frac{0.99c}{c}\right)^2}}$$

$$= \frac{1.77\times10^{-8}}{\sqrt{1-(0.99)^2}}$$

$= 1.3 \times 10^{-7}$ sec.

This is the half-life appropriate to the laboratory frame of reference. Pions that live this long, traveling at a speed of 0.99c, would cover a distance

$$d = 2.97 \times 10^8 \text{ m/sec} \times 1.3 \times 10^{-7}$$

$$= 39 \text{ m.}$$

which is the length measured in the laboratory.

Example 11.11 *Two events have the space time co-ordinates (0,0,0,0) and (4c, 0, 0, 3) in a given frame S. (i) What is the time-interval between them? (ii) Obtain the velocity of a frame in which (a) the two events occur simultaneously, and (b) the first event occurs two seconds earlier than the second.*

Soln.

Let

$$x_1 = 0 \qquad x_2 = 4c$$

$$y_1 = 0 \qquad y_2 = 0$$

$$z_1 = 0 \qquad z_2 = 0$$

$$t_1 = 0 \qquad t_2 = 3$$

(i) The time interval between these two co-ordinates

$$t = t_2 - t_1 = 3 \text{ seconds}$$

(ii) Let a frame of reference S' be moving with a velocity v relative to S along the x-axis. Let the two events occur at time t'_1 and t'_2 respectively in S'. Then

$$t'_1 = \frac{t_1 - \frac{vx_1}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{and} \quad t'_2 = \frac{t_2 - \frac{vx_2}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

(a) For the two events to occur simultaneously $t'_2 - t'_1 = 0$.

$$\text{or,} \quad 0 = \frac{(t_2 - t_1) - \frac{v(x_2 - x_1)}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{or,}\quad 3-\frac{v.4c}{c^2}=0$$

$$\text{or,}\quad \frac{4v}{c}=3$$

$$\text{or,}\quad v=\frac{3}{4}\times c=\frac{3}{4}\times 3\times 10^8=2.25\times 10^8\ \text{m/sec.}$$

(b) The first event occur two seconds earlier hence

$$t'_2-t'_1=-2=\frac{(t_2-t_1)-\frac{v(x_2-x_1)}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{or,}\quad -2=\frac{3-\frac{v.4c}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{or,}\quad 4\left(1-\frac{v^2}{c^2}\right)=\left(3-\frac{4v}{c}\right)^2$$

$$\text{or,}\quad 4(c^2-v^2)=(3c-4v)^2$$

$$\text{or,}\quad 4c^2-4v^2=9c^2-24vc+16v^2$$

$$\text{or,}\quad 20v^2-24vc+5c^2=0$$

$$\text{or,}\quad v=\frac{24c\pm\sqrt{(24c)^2-4\times 20\times 5}}{2\times 20}$$

$$\text{or,}\quad v=0.93c\quad \text{or,}\quad 0.268c.$$

$v = 0.268$ c gives a positive value for $t'_2 - t'_1$. Hence this cannot be the required value of v.

$$\therefore\quad v=0.94c=0.94\times 3\times 10^8=2.82\times 10^8\ \text{m/sec.}$$

Example 11.12 The length of a spaceship is measured to be exactly half its actual length. Calculate (i) the speed of the spaceship and (ii) the time dilation corresponding to one second on the spaceship.

Soln.

(i) $L = L_O\sqrt{1-\frac{v^2}{c^2}}$

$\frac{L}{L_O} = \sqrt{1-\frac{v^2}{c^2}}$ Here $\frac{L}{L_O} = 0.5$

$\therefore\ 0.5 = \sqrt{1-\frac{v^2}{c^2}}$ $c = 3 \times 10^8$ m/sec; v = ?

or, $\frac{v^2}{c^2} = 0.75$

$\therefore\ v = \sqrt{0.75}\,.\,c$

$= 0.866\ c$

$= 0.866 \times 3 \times 10^8$ m/sec

$= 2.598 \times 10^8$ m/sec.

(ii) The time t as observed from the stationary frame corresponding to the time $t_0 = 1$ second on the spaceship is given by

$$t_0 = t\sqrt{1-\frac{v^2}{c^2}}$$

or, $t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}}$ Here, $t_0 = 1$ second

$\frac{v^2}{c^2} = 0.75$

$\therefore\ = \frac{1}{\sqrt{1-0.75}}$ sec

$= 2$ sec.

Example 11.13 The proper length of a rod is 5 metres. What would be its length for an observer if it be moving relative to him in a direction parallel to its own length (i) with velocity 0.8c, (ii) moving with velocity c? What would be its length for an observer who is himself moving along with it?

Soln.

$$l = l_0\sqrt{1-\frac{v^2}{c^2}}$$ where l_0 is the proper length of the rod.

(i) when $v = 0.8\ c$

$$l = l_0\sqrt{1-\frac{(0.8c)^2}{c^2}}$$

$$= 5\sqrt{1-0.64}$$

$$= 5\sqrt{0.36}$$

$$= 5 \times 0.6 = 3\text{m}.$$

(ii) when $v = c$

$$l = l_0\sqrt{1-\frac{v^2}{c^2}} = 5 \times \sqrt{1-1} = 5 \times 0 = 0\ \text{m}.$$

when the observer moves with the rod, $v = 0$

$$\therefore \quad l = l_0\sqrt{1-\frac{v^2}{c^2}} = l_0\sqrt{1-0} = 5 \times 1 = 5\ \text{m}.$$

Example 11.14 How fast would a rocket ship have to go relative to an observer for its length to be contracted to 99% of its length when at rest?

Soln.

Let l_0 = proper length

then $\quad l = 0.99\ l_0$

Therefore $\frac{l}{l_0} = \frac{0.99\, l_0}{l_0} = \frac{99}{100}$

From, $l = l_0 \sqrt{1 - v^2/c^2}$, we have

$$\frac{l}{l_0} = \frac{99}{100} = \sqrt{1 - \frac{v^2}{c^2}}$$

or, $1 - \frac{v^2}{c^2} = \left(\frac{99}{100}\right)^2$

or, $\frac{v^2}{c^2} = 1 - \left(\frac{99}{100}\right)^2 = 0.0199$

or, $v^2 = 0.0199\, c^2$

or, $v = 0.1416\, c$

$= 0.1416 \times 3 \times 10^8$ m/sec

$= 4.245 \times 10^7$ m/sec.

Example 11.15 Obtain the volume of a cube, the proper length of each edge of which is l_0, when it is moving with a velocity v along one of its edges.

Soln.

Obviously, the only change in length, (*i.e.*, a contraction) will occur in the particular edge of the cube along which it is moving. The lengths of the other edges, being perpendicular to the direction of motion, will remain unaffected.

If l_0 be the proper length of each edge of the cube, the length of the edge along which it is moving will become $l_0\sqrt{1 - v^2/c^2}$.

Since the length of the other two edges remain unaffected, the length of those edges will remain same as before, *i.e.*, l_0.

Therefore, the volume of the moving cube.

$$= l_0\sqrt{1 - v^2/c^2} \cdot l_0 \cdot l_0.$$

$$= l_0^3 \sqrt{1-v^2/c^2}$$

***Example 11.16** Calculate the percentage contraction in the length of a rod in a frame of reference, moving with velocity 0.8c in a direction (a) parallel to its length (b) at an angle 30° with its length. What is the orientation of the rod in the moving frame of reference in case (b)?*

Soln.

(a) Let l_0 be the length of the rod placed along the x-axis in a frame of reference S. Then its length as measured in a frame of reference S' moving with a velocity 0.8 c in a direction parallel to the length of the rod *i.e.*, x-axis, is given by

$$l = l_0\sqrt{1-v^2/c^2} = l_0\sqrt{1-\frac{(0.8c)^2}{c^2}}$$

$$= l_0\sqrt{1-0.64} = l_0\sqrt{0.36} = 0.6 \times l_0.$$

∴ percentage contraction in the length

$$= \frac{l_0 - 0.6\, l_0}{l_0} \times 100 = 40\%$$

(b) In this case, the component of the length of the rod along its direction of motion

$$l_{11} = l_0 \cos 30° = \frac{\sqrt{3}}{2} l_0$$

and the component in a direction perpendicular to the direction of motion

$$l_{\perp} = l_0 \sin 30° = 0.5\, l_0$$

Only the component parallel to the direction of motion will suffer contraction and its length as measured in S' will be

$$= \frac{\sqrt{3}}{2} l_0 \sqrt{1-v^2/c^2}$$

$$= \frac{\sqrt{3}}{2} \times 0.6\, l_0 = 0.52\, l_0$$

As the component perpendicular to the direction of motion does not suffer any contraction, its length as measured in S' remain $0.5\, l_0$.

Therefore, the total length of the rod as measured in S'

$$= \sqrt{(0.52\, l_0)^2 + (0.5\, l_0)^2}$$

$$= 0.7228\, l_0$$

Hence the percentage contraction produced in the length of the rod

$$= \frac{l_0 - 0.7228\, l_0}{l_0} \times 100$$

$$= 22.72\%$$

If θ be the angle the rod appears to make with the direction of velocity as observed from S', then

$$\tan\theta = \frac{\text{length component} \perp}{\text{length component} \parallel}$$

$$= \frac{0.5\, l_0}{0.52\, l_0} = 096.$$

$$\theta = \tan^{-1}(0.96) = 43° \; 50'$$

The rod therefore makes an angle of 43°50' with its direction of motion.

Example 11.17 *The half-life of a particular particle, as measured in the laboratory comes out to be 4×10^{-8} sec, when its speed is 0.8 c and 3×10^{-8} sec, when its speed is 0.6 c. Explain*

Soln.

Let t_0 be the proper time and t the measured time.

$$\text{Then} \quad t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

In the first case

$$t_1 = 4 \times 10^{-8} \text{ sec} \qquad v = 0.8c.$$

$$\therefore \quad t_0 = t_1\sqrt{1-\frac{v^2}{c^2}} = 4 \times 10^{-8} \cdot \sqrt{1-\frac{(0.8c)^2}{c^2}}$$

$$= 4 \times 10^{-8}\sqrt{1-0.64}$$

$$= 4 \times 10^{-8}\sqrt{0.36} = 0.6 \times 4 \times 10^{-8}$$

$$= 2.4 \times 10^{-8} \text{ sec.}$$

In the second case

$$t_2 = 3 \times 10^{-8} \text{ sec} \qquad v = 0.6 \text{ c.}$$

$$\therefore \quad t_0 = 3 \times 10^{-8}\sqrt{1-\frac{(0.6c)^2}{c^2}}$$

$$= 3 \times 10^{-8}\sqrt{1-0.36}$$

$$= 3 \times 10^{-8}\sqrt{0.64}$$

$$= 3 \times 10^{-8} \times 0.8 = 2.4 \times 10^{-8} \text{ sec.}$$

The proper half-life is same in both cases. The observed difference in the measured value is due to relativistic time dilation.

Example 11.18 A clock in a spaceship emits signals at intervals of 1 second as observed by an astronaut in the spaceship. If the spaceship travels with a speed of 3×10^7 ms^{-1}, what is the interval between successive signals as seen by an observer at the control centre on the ground?

Soln.

Here $t_0 = 1$ sec; $\qquad v = 3 \times 10^7$ ms^{-1}

$c = 3 \times 10^8$ ms^{-1} $\qquad t = ?$

$$t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{(3\times10^7)^2}{(3\times10^8)^2}}} = 1.005 \text{ sec}$$

***Example 11.19** A particle with a proper life time of 1 μs moves through the laboratory at 2.7 x 10^7 ms^{-1}. (a) What is its life time, as measured by the observers in the laboratory? (b) What will be the distance traversed by it before disintegrating?*

Soln.

Here $t_0 = 1\ \mu s = 10^{-6}$ sec; $\quad v = 2.7 \times 10^8\ ms^{-1}$; t = ?

$$t = \frac{t_0}{\sqrt{1-\dfrac{v^2}{c^2}}} = \frac{10^{-6}}{\sqrt{1-\dfrac{(2.7\times10^8)^2}{(3\times10^8)^2}}} = 2.3 \times 10^{-6} \text{ sec.}$$

The average distance moved by the particle before disintegration = (2.7 x 10^8) (2.3 x 10^{-6}) = 620 m.

***Example 11.20** Determine the time (as measured by a clock at rest on rocket) taken by a rocket to reach a distant star and return to earth with a constant velocity $v = \sqrt{0.999}\ c$ if the distance of the star is 4 light years. (A light year is defined as the distance traveled by a light beam in vacuum in one year).*

Soln.

The time taken by the rocket for the round trip, *i.e.*, to the star and back on earth, as measured by a clock on the earth is

$$t = \frac{(4+4)c}{\sqrt{0.9999}\,c} \text{ year}$$

$$\approx 8 \text{ years.}$$

The same time as measured by a clock on the rocket, *i.e.*, proper time

$$t_0 = t.\sqrt{1-v^2/c^2} = 8\sqrt{1-0.9999}$$

$$= 0.08 \text{ years.}$$

Thus the time for the round trip, as measured by a clock carried by the rocket, is 0.08 years.

11.7 The relativistic addition of velocities

Let a train move with a velocity v with respect to the ground. If a passenger on the train, move with a velocity u' with respect to the train, then according to classical physics, the passenger's velocity relative to the ground u is just the vector sum of the two velocities, *i.e.*,

$$u = u' + v$$

Let us consider, for the moment, the special case wherein all velocities are along the common x - x' direction of the two inertial frames S and S'. Let S be the ground frame and S' the frame of the train, whose speed relative to the ground is v. The passenger's speed in the S' frame is u' and his position on the train as time goes on can be described by

$$x' = u't'$$

What is the speed of the passenger as observed from the ground? Using Lorentz transformation equations we have,

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} = u't'$$

and

$$t' = \frac{t - (v/c^2)x}{\sqrt{1 - v^2/c^2}}$$

Combining these results we have

$$x - vt = u'\left(t - \frac{v}{c^2}x\right)$$

which can be written as

$$x = \frac{(u' + v)}{1 + \frac{u'v}{c^2}} \cdot t \qquad (11.11)$$

If the passenger's speed relative to the ground be u, then his ground location as time goes on is given by

$$x = ut \qquad (11.12)$$

comparing eqn. (11.11) with eqn. (11.12), we obtain

$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}} \qquad (11.13)$$

This is the relativistic addition of velocities.

If u' and v are very small compared to c, then the second term in the denominator of eqn. (11.13) is negligible compared to one. Then eqn. (11.13) reduces to the classical result, *i.e.*, u = u' + v. If, on the other hand, u' = c *i.e.*, our passenger on the train is a light pulse, then

$$u = \frac{c + v}{1 + \frac{cv}{c^2}} = \frac{(c+v)}{c(c+v)}.c^2 = c$$

11.8 Relativity of mass

According to relativity, the measurements of an object or an event will be different for different observers who are in relative motion with each other. The mass m of a body moving with a velocity v is found to be given by a relation

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

where m_0 is the mass of the body when at rest and c is the velocity of light. Thus an object will appear to have more mass while in motion then when it is at rest.

Derivation :

Consider two frames of reference S and S'. S' is moving with a constant velocity v relative to S, in the positive x-direction. Suppose that two exactly similar elastic balls A and B in the frame S' approach each other at equal speed, say u and -u. Let the mass of each ball be m in S'. The two balls collide with each other and alter collision coalesce into one body. According to the law of conservation of momentum

momentum of ball A + momentum of ball B = momentum of coalesced mass

or, $mu + m(-u)$ = momentum of coalesced mass

$$= 0$$

Thus the coalesced mass must be at rest in S' frame.

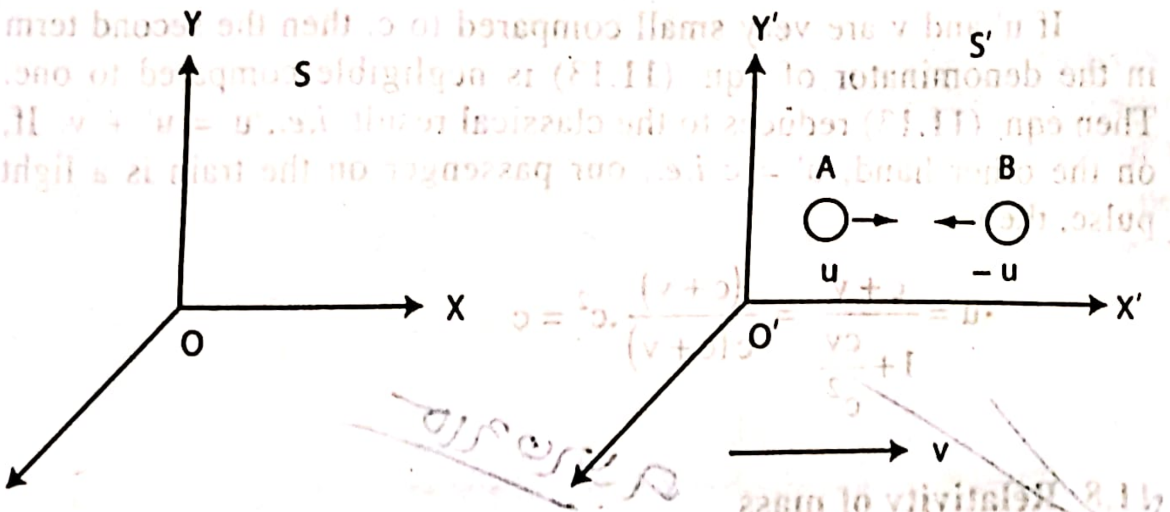


Fig. 11.4

Let the collision be considered with reference to the frame S. Let u_1 and u_2 be the velocities of the balls as observed from S. Then

$$u_1 = \frac{u+v}{1+\frac{uv}{c^2}} \qquad \text{.........................} \quad \text{(i)}$$

and $$u_2 = \frac{-u+v}{1-\frac{uv}{c^2}} \qquad \text{.........................} \quad \text{(ii)}$$

After collision, velocity of the coalesced mass is v relative to the frame S. As observed from the frame S, let the mass of the ball traveling with velocity u_1 be m_1 and that of B with velocity u_2 be m_2. Since the total momentum of the balls must be conserved, we have

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v \text{ ....} \quad \text{(iii)}$$

substituting the value of u_1 and u_2 as given by equations (i) and (ii) in equation (iii), we get

$$m_1\left[\frac{u+v}{1+\frac{uv}{c^2}}\right]+m_2\left[\frac{-u+v}{1-\frac{uv}{c^2}}\right]=(m_1+m_2)v$$

$$\text{or,}\quad m_1\left[\frac{u+v}{1+\frac{uv}{c^2}}-v\right]=m_2\left[v-\frac{-u+v}{1-\frac{uv}{c^2}}\right]$$

$$\text{or,}\quad m_1\left[\frac{u+v-v-\frac{uv^2}{c^2}}{1+\frac{uv}{c^2}}\right]=m_2\left[\frac{v-\frac{uv^2}{c^2}+u-v}{1-\frac{uv}{c^2}}\right]$$

$$\text{or,}\quad m_1\left[\frac{u\left(1-\frac{v^2}{c^2}\right)}{1+\frac{uv}{c^2}}\right]=m_2\left[\frac{u\left(1-\frac{v^2}{c^2}\right)}{1-\frac{uv}{c^2}}\right]$$

$$\text{or,}\quad \frac{m_1}{m_2}=\frac{1+\frac{uv}{c^2}}{1-\frac{uv}{c^2}} \qquad \text{.......................} \qquad \text{(iv)}$$

$$\text{Also,}\quad 1-\frac{u_1^2}{c^2}=1-\frac{\left(\frac{u+v}{c}\right)^2}{\left(1+\frac{uv}{c^2}\right)^2}$$

$$=\frac{1+\frac{u^2v^2}{c^4}+\frac{2uv}{c^2}-\frac{u^2}{c^2}-\frac{v^2}{c^2}-\frac{2uv}{c^2}}{\left(1+\frac{uv}{c^2}\right)^2}$$

$$= \frac{\left(1-\frac{u^2}{c^2}\right)-\frac{v^2}{c^2}\left(1-\frac{u^2}{c^2}\right)}{\left(1+\frac{uv}{c^2}\right)^2}$$

$$\therefore \quad 1-\frac{u_1^2}{c^2} = \frac{\left(1-\frac{u^2}{c^2}\right)\left(1-\frac{v^2}{c^2}\right)}{\left(1+\frac{uv}{c^2}\right)^2} \quad \text{.................(v)}$$

Similarly, $1-\frac{u_2^2}{c^2} = \frac{\left(1-\frac{u^2}{c^2}\right)\left(1-\frac{v^2}{c^2}\right)}{\left(1-\frac{uv}{c^2}\right)^2}$ .... (vi)

Dividing equation (vi) by (v)

$$\frac{1-\frac{u_2^2}{c^2}}{1-\frac{u_1^2}{c^2}} = \frac{\left(1+\frac{uv}{c^2}\right)^2}{\left(1-\frac{uv}{c^2}\right)^2}$$

or, $$\frac{\sqrt{1-\frac{u_2^2}{c^2}}}{\sqrt{1-\frac{u_1^2}{c^2}}} = \frac{1+\frac{uv}{c^2}}{1-\frac{uv}{c^2}} \quad \text{............... (vii)}$$

From equation (vii) and (iv), we get

$$\frac{m_1}{m_2} = \frac{\sqrt{1-\frac{u_2^2}{c^2}}}{\sqrt{1-\frac{u_1^2}{c^2}}}$$

or, $$m_1\sqrt{1-\frac{u_1^2}{c^2}} = m_2\sqrt{1-\frac{u_2^2}{c^2}} \quad \text{.............. (viii)}$$

Since the L.H.S. and R.H.S. of eqn. (viii) are independent of one another, the above result can be true only if each is a constant.

Therefore,

$$m_1\sqrt{1-\frac{u_1^2}{c^2}} = m_2\sqrt{1-\frac{u_2^2}{c^2}} = m_0 \quad \text{......... (ix)}$$

The constant denoted by m_0, is called the rest mass of the body and corresponds to *zero velocity*.

Thus $m_1 = \dfrac{m_0}{\sqrt{1-\frac{u_1^2}{c^2}}}$; or $m_2 = \dfrac{m_0}{\sqrt{1-\frac{u_2^2}{c^2}}}$

In general, if m denotes the mass of a body when it is moving with a velocity v, then,

$$m_1 = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \quad (11.14)$$

Equation (11.14) is the relativistic formula for the variation of mass with velocity. If v → c in equation (11.14), m → ∝ *i.e.*, an object traveling at the speed of light would have infinite mass. Thus it is clear from equation (11.14) that no material body can have a velocity equal to or greater than the velocity of light.

11.9 Relationship between the total energy, the rest energy and the momentum

we have $$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{m_0}{\left(1-\frac{v^2}{c^2}\right)^{1/2}}$$

Expansion by the binomial theorem yields,

$$m = m_0 \left(1 + \frac{v^2}{2c^2} + \frac{3}{8}\frac{v^4}{c^4} + \ldots\ldots\ldots\ldots\right)$$

Rearranging, the above expression may be rewritten as

$$mc^2 = m_0c^2 + \frac{1}{2}\,m_0v^2 + \ldots\ldots\ldots\ldots$$

The terms $\frac{1}{2}\, m_0v^2$ is already familiar to us and represents the kinetic energy of a body of mass m_0, moving with a velocity v. It can be seen that mc^2 is the total energy of a body moving with velocity v, m_0c^2 is the energy associated with a mass m_0 at rest. Thus it appears that mass is to be regarded as a form of energy with c^2 as the conversion factor. The complete expression for the total energy E as given in art. 11.10

$$E = mc^2 = \frac{m_0c^2}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

and the momentum

$$p = mv = \frac{m_0 v}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

From eqn. (i), we have

$$\left(\frac{E}{c}\right)^2 = \frac{m_0^2c^2}{\left(1 - \frac{v^2}{c^2}\right)}$$

$$\text{or,}\quad m_0^2c^2 = \left(\frac{E}{c}\right)^2 - \left(\frac{E}{c}\right)^2 \cdot \frac{v^2}{c^2}$$

Using eqn. (ii),

$$m_0^2c^2 = \left(\frac{E}{c}\right)^2 - \left(\frac{mc^2}{c}\right)^2 \cdot \frac{v^2}{c^2}$$

$$= \left(\frac{E}{c}\right)^2 - m^2v^2 = \left(\frac{E}{c}\right)^2 - p^2$$

$$\therefore \quad E^2 = m_0^2c^4 + p^2c^2. \tag{11.15}$$

Eqn. (11.15) gives the relationship between the total energy, the rest energy and the momentum. Thus for a mass at rest p = 0 and $E = m_0c^2$. A particle with no rest mass (m_0 = 0) can still have a momentum given by $p = \frac{E}{c} = mc$ as in the case of photon in the Compton effect.

Now, $E = m_0c^2 + E_K$. Hence from eqn. (11.15)

$$p = \sqrt{2m_0E_K + \frac{E_K^2}{c^2}} \tag{11.16}$$

Except for the second term under the square root, eqn. (11.16) is identical with the classical formula for momentum. The last term is therefore called the *relativistic - correction term.*

11.10 Mass and Energy

The most famous relationship Einstein obtained from the postulates of special relativity concerns mass and energy. When a steady net force is applied to a material body of rest mass m_0, the body increases in speed. Since the force is acting through a distance, work is done on the body and its energy increases. On the other hand, the mass of the body increases with increasing speed. Thus the work done on a material body not only increases its speed but also contributes to increasing its mass. Normally the work done on a body increases its energy. This new twist from the theory of relativity leads to the idea that mass is a form of energy, a crucial part of Einstein's theory of relativity.

To find the mathematical relationship between mass and energy, it will be assumed that

(i) the definition of force is given by Newton's second law

(ii) the classical definition of work and the relationship between work and the change in kinetic energy (work-energy theorem) are still valid in relativity.

In classical physics, force is defined as the time rate of change of momentum. Or

$$F = \frac{d}{dt}(mv) = m\frac{dv}{dt} \quad \text{.................. (i)}$$

since mass is constant. This definition of force is kept in relativistic mechanics, but now mass is also a variable. Hence

$$F = \frac{d}{dt}(mv) = m\frac{dv}{dt} + v\frac{dm}{dt} \quad \text{........... (ii)}$$

Suppose the kinetic energy of a body is increased by an amount dT by exerting a force through a distance. Then the change in kinetic energy can be found, as in classical physics, by calculating the work done by the force. Or,

$$dT = F \cdot dx \quad \text{.................. (iii)}$$

where F is the force acting on the body and dx is the distance through which it works.

Substituting eqn. (ii) in eqn. (iii), we get

$$dT = m.\frac{dv}{dt}.dx + v.\frac{dm}{dt}.dx \quad \text{......... (iv)}$$

Since $dx/dt = v$, eqn. (iv) can be written as

$$dT = mvdv + v^2dm \quad \text{.................. (v)}$$

$$= v\,(mdv + vdm) = v.d(mv) \quad \text{...... (v.a)}$$

Eqn. (v) can be simplified by using mass - transformation equation. Differentiating both sides of equation

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

we get

$$dm = \frac{m_0}{c^2} \frac{vdv}{\left[1-\frac{v^2}{c^2}\right]^{3/2}}$$

$$= \frac{mvdv}{c^2 - v^2} \qquad \text{.................. (vi)}$$

or, $mvdv = (c^2 - v^2)\, dm \qquad$ .... (vii)

Substituting the value of mvdv as given by eqn. (vii) in eqn. (v), we get

$$dT = (c^2 - v^2)\, dm + v^2 dm$$

$$= c^2 dm \qquad \text{.................. (viii)}$$

Eqn. (viii) shows that in relativity a change in kinetic energy, dT, is directly proportional to a change in mass, dm, the proportionality factor being the square of the velocity of light. The integral form of this equation can also be obtained easily. The increase in kinetic energy as given by eqn. (v.a) is

$$dT = vd\,(mv)$$

Since the body starts from rest (*i.e.*, v = 0) and finally acquires a velocity v, we have

Kinetic energy

$$T = \int_0^v vd\,(mv)$$

$$= \int_0^v vd\left(\frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}\right) \qquad \left(\because\ m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}\right)$$

Integrating by parts [*i.e.*, $\int xdy = xy - \int ydx$; we have]

$$T = \frac{m_0 v^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0 \int_0^v \frac{vdv}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \frac{m_0 v^2}{\sqrt{1-\frac{v^2}{c^2}}} + m_0 c^2 \left[\sqrt{1-\frac{v^2}{c^2}} \right]_0^v$$

$$= \frac{m_0 c^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0 c^2 \quad \text{.................... (ix.a)}$$

$$= mc^2 - m_0 c^2$$

$$= c^2 (m - m_0) \quad \text{.................... (x)}$$

[Alternately: when m_0 is the mass of the body when $v = 0$ and m is the mass of the body when $v = v$, we have

$$T = \int_0^v dT = c^2 \int_{m_0}^m dm = c^2 (m - m_0) \text{ .... (x.a)]}$$

Thus, the kinetic energy of a body at high speeds is not equal to $\frac{1}{2}mv^2$; nor can we substitute $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$ into the classical expression and get a correct result *i.e.*, K.E. $\neq \frac{\frac{1}{2}m_0 v^2}{\sqrt{1-\frac{v^2}{c^2}}}$

Eqn. (x) requires some interpretation. According to this equation, the kinetic energy of a body can be expressed in terms of the increase in the mass of the body. Consistent with the idea that mass is a form of energy, eqn. (x) may be further interpreted as meaning that the rest mass m_0 is associated with an amount of energy m_0c^2 *i.e.*, when the body is at rest, the internal energy stored in the body is m_0c^2. Einstein called m_0c^2 the *rest mass energy* (or proper energy) of the body. The total energy E possessed by the body is then the sum of the *rest mass energy* and *kinetic energy* of the body. Or,

$$E = T + m_0c^2 \qquad \text{(xi)}$$

$$= mc^2 - m_0c^2 + m_0c^2$$

$$= mc^2 \qquad (11.17)$$

$$= \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}}$$

Eqn. (11.17) gives Einstein's famous formula $E = mc^2$ and is referred to as Einstein's mass-energy relation.

Like the other transformation equations, the expression for the kinetic energy should reduce to the classical expression for $v << c$. From eqn. (xi), we have

$$T = E - m_0c^2 = mc^2 - m_0c^2$$

$$= \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0c^2$$

Now since $v << c$, we can say

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \left(1-\frac{v^2}{c^2}\right)^{-1/2}$$

$$= \left(1+\frac{1}{2}\cdot\frac{v^2}{c^2}+\frac{3}{8}\,\frac{v^4}{c^4}+\ldots\ldots\ldots\ldots\right)$$

where three terms of the binomial expansion have been carried out. We then have,

$$T = m_0c^2\left(1+\frac{1}{2}\cdot\frac{v^2}{c^2}+\frac{3}{8}\,\frac{v^4}{c^4}+\ldots\ldots\ldots\ldots\right) - m_0c^2$$

$$= m_0c^2\left(1+\frac{1}{2}.\frac{v^2}{c^2}+\frac{3}{8}\frac{v^4}{c^4}+..............-1\right)$$

$$= \frac{1}{2}m_0v^2+\frac{3}{8}m_0\frac{v^4}{c^4}+.................. \quad (11.18)$$

The first term in eqn. (11.18) is the classical expression for the kinetic energy. Obviously, it is the only significant term for low velocities.

Eqn. (11.17) suggests further, in plain and simple terms, that *mass may appear as energy and energy as mass*, the conversion factors being c^2 and $1/c^2$ respectively. A kilogram of matter has an energy content of 9 x 10^{16} joules. Even a minute bit of matter represents a vast amount of energy, and in fact, the conversion of matter into energy is the source of the power liberated in all of exothermic reactions of physics and chemistry. The fact that mass can be converted into energy and *vice versa*, suggests that in addition to kinetic, potential, electromagnetic, thermal and other familiar forms, energy can also manifest itself as mass. This calls for the widening of the scope of the law of conservation of energy. Since mass and energy are not independent entities, the separate conservation principles of energy and mass are properly a single one, the *principle of conservation of mass-energy*. It is neither mass alone nor energy alone that remains conserved but mass, inclusive of energy in terms of mass, or energy, that is really conserved. Mass can be created or destroyed, but only if an equivalent amount of energy simultaneously vanishes or comes into being and *vice versa.*

It is interesting to note that as v → c in eqn. (ix.a) the kinetic energy T tends to infinity. That is, from eqn. (x.a) an infinite amount of work would need to be done on the particle to accelerate it up to the speed of light. Once again we find c playing the role of a limiting velocity. It can also be noted that from the expression $T = (m - m_0)c^2$ that, a change in the kinetic energy of a particle is related to a change in its (inertial) mass.

11.11 Rest mass of a photon

According to eqn. 11.17, associated with mass m, there is an amount of energy mc^2; conversely, to an energy E, there corresponds a mass given by

$$m = \frac{E}{c^2}$$

Another important consequence of the proportionality between mass and energy has to do with the momentum associated with the transfer of energy. If a quantity of energy is transferred with a velocity v, then the magnitude of the associated momentum can be written as

$$p = mv = \frac{E}{c^2}.v \qquad (11.19)$$

The relationship as given by eqn. (11.19) can be applied, for example, to energy and momentum carried by a light quantum (photon). The energy carried by a photon is hν and its velocity is c. Then the momentum is, according to eqn. (11.19)

$$p = \frac{E}{c^2}.c = \frac{E}{c} = \frac{h\nu}{c} \qquad (11.20)$$

Eqn. (11.20) shows why the rest mass of a photon must be zero. For, if the rest mass is not zero, then the mass of a photon traveling with the velocity of light would be infinite, because $\sqrt{1-\frac{v^2}{c^2}}$ becomes zero for v = c. Although the rest mass of a photon is zero, the photon has a mass associated with its kinetic energy and, according to eqn. (11.17) this mass is $\frac{h\nu}{c^2}$. The momentum corresponding to this mass would then be $\frac{h\nu}{c}$.

We can now obtain an important relation between the total mass energy E of a body and its momentum p (= mv). Let us consider the mass-transformation equation

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

Squaring and then rearranging, the above expression leads to

$$m^2c^2 = m^2v^2 + m_0^2c^2$$

Multiplying both sides of this equation by c^2 and then making the appropriate substitutions from $E = mc^2$ and $p = mv$, we get

$$E^2 = p^2c^2 + m_0^2c^4$$

Again,

$$E = m_0c^2 + T$$

where T is the kinetic energy of the particle. By substituting E in eqn. (11.15) and solving for the momentum, we obtain

$$p = \sqrt{2m_0T + T^2/c^2}$$

Except for the second term under the square root, this equation is identical with the classical formula for the momentum. The last term is therefore called the *relativistic-correction term.*

***Example 11.21** A particle is moving with a speed of 0.5c. Calculate the ratio of its rest mass and the mass while in motion.*

Sole.

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

or, $$\frac{m_0}{m} = \sqrt{1 - \frac{v^2}{c^2}}$$

Here $v = 0.5c$

$$\therefore \quad \frac{m_0}{m} = \sqrt{1 - \frac{(0.5c)^2}{c^2}}$$

$$= \sqrt{1 - (0.5c)^2}$$

$$= 0.866.$$

***Example 11.22** Calculate the velocity that one atomic mass unit will have if it has a kinetic energy equal to twice the rest mass energy.*

Soln.

We have

$$E = mc^2 = m_0c^2 + T$$

Here $T = 2m_0c^2$

$\therefore \; mc^2 = m_0c^2 + 2m_0c^2$

or, $3m_0c^2 = mc^2$

or, $m = 3m_0$

But $\quad m = \dfrac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}$

$\therefore \quad 3m_0 = \dfrac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}$

or, $\sqrt{1-\dfrac{v^2}{c^2}} = \dfrac{m_0}{3m_0} = \dfrac{1}{3}$

or, $1-\dfrac{v^2}{c^2} = \dfrac{1}{9}$

or, $\dfrac{v^2}{c^2} = 1-\dfrac{1}{9} = \dfrac{8}{9}$

or, $\quad v^2 = \left(\dfrac{8}{9}\right)c^2$

or, $\quad v = \sqrt{\dfrac{8}{9}}\cdot c$

$= 0.941$ c.

Example 11.23 For what value of v/c (=β) will be relativistic mass of a particle exceed its rest mass by a given fraction f?

Soln.

From eqn. $m = \frac{m_0}{\sqrt{1-v^2/c^2}}$

we have

$$f = \frac{m - m_0}{m_0} = \frac{m}{m_0} - 1 = \frac{1}{\sqrt{1-\beta^2}} - 1$$

or, $$\frac{1}{\sqrt{1-\beta^2}} = 1 + f$$

Squaring

$$\frac{1}{1-\beta^2} = (1 + f)^2$$

or, $$1 - \beta^2 = \frac{1}{(1+f)^2}$$

or, $$\beta^2 = 1 - \frac{1}{(1+f)^2} = \frac{1+2f+f^2-1}{(1+f)^2} = \frac{2f+f^2}{(1+f)^2} = \frac{f(2+f)}{(1+f)^2}$$

or, $$\beta = \frac{\sqrt{f(2+f)}}{(1+f)}$$

The table below shows some computed values, which hold for all particles regardless of their rest mass.

f	β
0.001 (0.1 percent)	0.014
0.01	0.14
0.1	0.42
1 (100 percent)	0.87
10	0.994
100	0.999

***Example 22.24** What is the length of a metre stick moving parallel to its length when its mass is 3/2 of its rest mass?*

Soln.

We have $m = \dfrac{m_0}{\sqrt{1-v^2/c^2}}$ .... (i)

or, $\dfrac{m}{m_0} = \dfrac{1}{\sqrt{1-v^2/c^2}}$; or $\dfrac{m_0}{m} = \sqrt{1-v^2/c^2}$

and $L = L_0\sqrt{1-v^2/c^2}$ .... (ii)

Dividing (ii) by (i)

$$\frac{L}{m} = \frac{L_0}{m_0}\left(1-v^2/c^2\right)$$

$$\therefore\ L = \frac{m}{m_0}\cdot L_0\left(1-v^2/c^2\right) = \left(\frac{m}{m_0}\right)L_0\left(\frac{m_0}{m}\right)^2 = \frac{m_0}{m}\cdot L_0$$

Here $\dfrac{m_0}{m} = \dfrac{2}{3}$ and $L_0 = 1$ metre

$$\therefore\ L = \frac{2}{3} \times 1 = 0.667\text{m}.$$

***Example 11.25** A particle of mass 10^{-24} kg is moving with a speed of 1.8×10^8 m/s. Calculate its mass when it is motion.*

Soln.

$$m = \frac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}$$

Here $m_0 = 10^{-24}$ kg
$v = 1.8 \times 10^8$ m/sec
$c = 3 \times 10^8$ m/sec
$\dfrac{v}{c} = 0.6$

$$\therefore\ m = \frac{10^{-24}}{\sqrt{1-(0.6)^2}}$$

$$= 1.25 \times 10^{-24}\ \text{kg}.$$

Example 11.26 *4.18 x 10^{-3} kg of a substance is fully converted to heat energy. Calculate the amount of heat produced.*

Soln.

Here $m = 4.18 \times 10^{-3}$ kg

$C = 3 \times 10^8$ m/sec.

$$E = mc^2$$

$$= 4.18 \times 10^{-3} \times (3 \times 10^8)^2 \text{ J}$$

$$\therefore \text{ Heat produced} = \frac{E}{4.18} \text{ calories}$$

$$= \frac{4.18 \times 10^{-3} \times (3 \times 10^8)^2}{4.18}$$

$$= 9 \times 10^{13} \text{ calories.}$$

Example 11.27 *Two particles each of rest mass 3 x 10^{-25} kg approach each other in head on collision. If each particle has an initial velocity of 2 x 10^8 m/sec, calculate the velocity of one particle as seen by the other. What is the relativistic mass of one seen by the other?*

Soln.

The velocity of one particle as seen by the other is given by

Here $u = 2 \times 10^8$ m/sec

$v = -2 \times 10^8$ m/sec

$c = 3 \times 10^8$ m/sec

$$u' = \frac{u - v}{1 - \frac{uv}{c^2}}$$

$$= \frac{4 \times 10^8}{1 + \left(\frac{4 \times 10^{16}}{9 \times 10^{16}}\right)}$$

$$= \frac{36}{13} \times 10^8$$

$$= 2.769 \times 10^8 \text{ m/sec.}$$

The relativistic mass of one particle as seen by the other is given by

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

Here $m_0 = 3 \times 10^{-25}$ kg

$c = 3 \times 10^8$ m/sec

$v = u' = 2.769 \times 10^8$ m/sec

$$= \frac{3\times10^{-25}}{\sqrt{1-\left(\frac{2.769\times10^8}{3\times10^8}\right)^2}}$$

$$= \frac{3\times10^{-25}}{\sqrt{1-\left(\frac{12}{13}\right)^2}}$$

$= 7.8 \times 10^{-25}$ kg.

***Example 11.28** Electrons are accelerated upto a kinetic energy of 10^9 eV. Find (i) the ratio of their mass to the rest mass (ii) the ratio of their velocity to the velocity of light and (iii) the ratio of their energy to the rest mass energy.*

Soln.

$U = 10^9$ eV $= 10^9 \times 1.6 \times 10^{-19}$ J

Also $U = mc^2$

$$m = \frac{U}{c^2} = \frac{10^9 \times 1.6\times10^{-19}\,\text{J}}{(3\times10^8)^2}$$

$= 1.77 \times 10^{-27}$ kg.

Rest mass of electron,

$m_0 = 9 \times 10^{-31}$ kg.

(i) $\frac{m}{m_0} = \frac{1.77\times10^{-27}}{9\times10^{-31}} = 1.95 \times 10^3$

(ii) $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$

$$\text{or, } \frac{m}{m_0} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$1.95 \times 10^3 = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$1-\frac{v^2}{c^2} = \left[\frac{1}{1.95\times10^3}\right]^2$$

$$\frac{v^2}{c^2} = 1 - \left[\frac{1}{1.95\times10^3}\right]^2$$

$$\frac{v}{c} = \left[1-\left\{\frac{1}{1.95\times10^3}\right\}^2\right]^{½}$$

$$= \left[1-\frac{1}{2\times(1.95)^2\times10^6}\right]^2$$

$$= \left[1-1.315\times10^{-7}\right]$$

(iii) Rest mass energy,

$$U_0 = m_0c^2 = 9 \times 10^{-31} \times (3 \times 10^8)^2$$

$$= 8.1 \times 10^{-14} \text{ J}$$

$$\frac{U}{U_0} = \frac{10^9\times1.6\times10^{-19}}{8.1\times10^{-14}} = 1.975 \times 10^3.$$

Example 11.29 *At what speed will an electron move in order to double its relative mass to the rest mass?*

Soln.

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = 2m_0$$

or, $$\frac{m_0}{m} = \frac{1}{2} = \sqrt{1-\frac{v^2}{c^2}}$$

or, $$1 - \frac{v^2}{c^2} = \frac{1}{4}$$

or, $$v^2 = 0.75\ c^2$$

$$v = 0.866\ c$$

$$= 0.866 \times 3 \times 10^8$$

$$= 2.598 \times 10^8 \text{ m/sec.}$$

Example 11.30 An electron has a total energy of 2 MeV. Calculate the effective mass of the electron in kg and also its speed. Assume rest mass of the electron to be 0.511 MeV.

Soln.

Here $$U = 2 \text{ MeV} = 2 \times 10^6 \times 1.6 \times 10^{-19} \text{ J}$$

$$= 3.2 \times 10^{-13} \text{ J}$$

$$U = mc^2$$

$$\therefore\ m = \frac{U}{c^2} = \frac{3.2\times10^{-13}}{(3\times10^8)^2} = 35.6 \times 10^{-31} \text{ kg.}$$

(ii) rest mass, $$m_0 = \frac{U}{c^2}$$

$$= \frac{0.511\times10^6\times1.6\times10^{-19}}{(3\times10^8)^2}$$

$$= 9.1 \times 10^{-31} \text{ kg.}$$

(iii) $m = \dfrac{m_0}{\sqrt{1-\dfrac{v^2}{c^2}}}$

or, $\dfrac{m}{m_0} = \dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}$

or, $\sqrt{1-\dfrac{v^2}{c^2}} = \dfrac{m_0}{m} = \dfrac{9.1\times10^{-31}}{35.6\times10^{-31}} = 0.2556$

$1-\dfrac{v^2}{c^2} = (0.2556)^2 = 0.06533$

or, $\dfrac{v^2}{c^2} = 1 - 0.06533 = 0.93467$

or, $v^2 = 0.93467c^2$

or, $v = 0.96$ c.

Example 11.31 Show that the rest mass of an electron (9.11×10^{-28} (gm) is equivalent to 0.511 MeV.

Soln.

The rest mass energy of an electron

$= m_0c^2$

$= 9.11 \times 10^{-28} \times 10^{-3} \times (3 \times 10^8)^2$ joules

$= 9.11 \times 10^{-31} \times 9 \times 10^{16}$ joules

$= 9.11 \times 9 \times 10^{-15}$ joules

$= \dfrac{9.11\times9\times10^{-15}}{1.6\times10^{-19}}$ eV $\quad$ ($1\text{eV} = 1.6 \times 10^{-19}$ J)

$= \dfrac{9.11\times9\times10^{-15}}{1.6\times10^{-19}\times10^6}$ MeV

$= 0.511$ MeV.

***Example 11.32** Calculate the kinetic energy of an electron with a velocity of 0.98 c times the velocity of light in the laboratory system.*

Soln.

When the electron moves with a velocity 0.98 c, its mass becomes

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{where } m_0 \text{ is the rest mass of the electron.}$$

$$= \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = 5.02\ m_0$$

Relativistic formula for kinetic energy

$$T = (m - m_0)\,c^2 = (5.02\ m_0 - m_0)\,c^2$$

$$= 4.02\ m_0 \times c^2 = 4.02 \times 9.11 \times 10^{-31} \times (3 \times 10^8)^2\ \text{J}$$

$$= 3.396 \times 10^{-13}\ \text{J}.$$

***Example 11.33** Show that the momentum of a particle of rest mass m_0 and kinetic energy K_E is given by the expression.*

$$p = \sqrt{\frac{K_E^2}{c^2} + 2m_0K_E}$$

Soln.

We have the relation $E^2 - p^2c^2 = m_0^2c^4$

$$\text{whence} \quad E^2 = m_0^2c^4 + p^2c^2$$

$$\text{or,} \quad E = (m_0^2c^4 + p^2c^2)^{1/2}$$

Again, the total energy E = rest energy + kinetic energy

$$\text{or,} \quad E = m_0c^2 + K_E$$

Equating the two values of E, we have

$$(m_0^2c^4 + p^2c^2)^{1/2} = m_0c^2 + K_E$$

Squaring both sides of the equation, we have

$$m_0^2c^4 + p^2c^2 = m_0^2c^4 + 2m_0c^2K_E + K_E^2$$

or, $p^2c^2 = K_E^2 + 2m_0c^2K_E$

or, $p^2 = \frac{K_E^2}{c^2} + 2m_0K_E$

$\therefore \quad p = \sqrt{\frac{K_E^2}{c^2} + 2m_0K_E}$, the required relation.

***Example 11.34** Calculate the mass of the electron when it is moving with a K.E. of 10 MeV.*

Soln.

K. E. $= (m - m_0)\, c^2 = 10$ MeV $= 10 \times 1.6 \times 10^{-13}$ J

$$\therefore \quad m = m_0 + \frac{10\times1.6\times10^{-13}\text{J}}{(3\times10^8)^2}$$

$= 9.1 \times 10^{-31} + 176 \times 10^{-31}$

$= 185.1 \times 10^{-31}$ Kg.

***Example 11.35** Does the mass of a substance increase on melting? Why?*

Soln.

Yes. Because an amount of energy equal to the specific latent heat of fusion has been added to the substance.

EXERCISES

1. Describe the Michelson-Morley experiment and explain the physical significance of the negative result.

2. State and explain the basic postulates of Einstein's special theory of relativity. Derive the Lorentz space time transformation formulae. Show that for values of $v << c$, Lorentz transformation reduces to the Galilean one.

3. Describe the Michelson-Morley experiment. Discuss the negative result of the experiment and show how it led to the special theory of relativity.

4. What do you understand by length contraction and time dilation? What are proper length and proper interval of time?

5. State the basic postulates of special theory of relativity and hence obtain the Lorentz space-time transformation formulae.

6. Obtain Lorentz space-time transformation formulae and hence obtain expressions for (i) length contraction and (ii) time dilation.

7. Deduce the formula for the relativistic variation of mass with velocity. Briefly explain its significance.

8. What do you understand by the relativity of mass? Show that the mass of a body becomes infinite when it moves with the velocity of light. What do you understand by this result?

9. What is the meaning of mass-energy relation? Derive Einstein's mass-energy relation.

10. Write down the relativistic expression for kinetic energy of a body and show that for smaller speeds it reduces to the classical expression.

11. State and deduce the mathematical expression for the law of addition of relativistic velocities. Show that in no case can the resultant velocity of a material particle be greater than c.

12. Describe Michelson-Morley experiment. With what aim in view was this experiment performed? How did Einstein explain the negative result of this experiment?

13. Derive an equation for the relativistic Kinetic energy of a body. How does this equation lead to the equivalence of mass and energy?

14. Write short explanatory notes, from the relativistic stand point, on (i) mass-energy equivalence, (ii) ultimate speed of a material particle, (iii) binding energy, (iv) pair production, (v) electron positron annihilation.

15. *Calculate the expected fringe shift in the Michelson-Morley experiment if the effective length of each path be 6 metres, velocity of earth, 3×10^4 m/sec and the wavelength of monochromatic light used, 5000 Angstrom units.*

Ans. 0.24 fringe-width.

16. *Show that a four dimensional volume element dx dy dz dt is invariant to Lorentz transformation.*

17. *The space-time coordinates of two events in frame S are 0,0,0,0 and 5c, 0, 0, 4 respectively. Obtain the space-time interval between them. What should be the velocity of a frame S' relative to S, in which (i) the two events may appear to occur simultaneously, (ii) the second may appear to occur 2 seconds (a) earlier, (b) later, than the first event?*

Ans: Space-time interval 3c. (1) 0.80 c (ii) (a) 0.44 c and (b) 0.94c.

18. *(a) A reference frame P' moves with respect to another frame P with a uniform velocity v. Write down the transformations giving x', y', z', t' in terms of x, y, z, t in Lorentz form. (The frames coincide at t = 0).*

(b) Show that for values v << c, Lorentz transformation reduces to the Galilean one.

19. *Show that the relativistic equations to transform space and time to system S from observations on system S' are $x = k(x' + vt')$ and $t = k[t' + vx'/c^2]$ respectively, where S' is moving relative to S with velocity v along the axis of x . $k = 1\sqrt{1-v^2/c^2}$.*

20. **A rod 1 metre long is moving along its length with a velocity 0.6 c. Calculate its length as it appears to an observer (a) on the earth (b) moving with the rod itself.**

Ans. (a) 0.8m (b) 1 m.

21. **Calculate the percentage contraction of a rod moving with a velocity 0.8c in a direction inclined at 60° to its own length.**

Ans. 9%.

22. **What is meant by length contraction due to relativistic effect? Show that length along the x-axis in frame S is contracted for an observer in frame S' and a length along the x'-axis in frame S' is contracted for an observer in frame S, where frame S' is moving with uniform velocity v with respect to S along the axis of x.**

What is meant by proper length.

23. The length of the side of a square, as measured by an observer in a stationary frame of reference S, is l. What will be its apparent area, as observed by him in a reference frame S' moving relative to S with velocity v along one of the sides of the square?

Ans. $l^2\sqrt{1-v^2/c^2}$.

24. A rod of true length 100 cm is moving with velocity 0.6c in a direction making an angle of 30° with its length. What is the contraction produced in its length and along what direction does it appear to move?

Ans. 15 cm, at an angle $\theta = \tan^{-1}$ (0.72) with the direction of velocity.

25. Show that if the time interval between two events occurring at the same place on the S' system be Δt', then the time interval observed in the S system is $\Delta t = k\Delta t'$ where $k = 1\sqrt{1-v^2/c^2}$. Is the time interval measured in S dilated or contracted with respect to that in S'?

Ans. Dilated.

26. A rocket ship travels away from the earth at 300 ms^{-1}. How many years must elapse before a clock in the ship and one on the ground differ by one second?

Ans. 6.3×10^4 years.

27. How fast should a rocket ship move relative to an observer in order that one year on it may correspond to two years on the earth'?

Ans. 2.598×10^8 ms^{-1}.

28. A π meson has a mean lifetime of 2×10^{-8} s when measured at rest. How far does it go before decaying into another particle if its speed is 0.99c?

29. A burst of 10^4 π^+ mesons, whose proper mean life is 2.5×10^{-8} sec, cover a circular path of radius 20 metres at a speed of 0.99c. (a) How many survive when the burst returns to the point of origin? (b) How many would be left in a burst that had remained at rest at the origin for the same period of time?

Ans. (a) 925 (b) none.

30. A group of muons, with a half-life of 1.5 μ sec, are produced at a height of 30 Km above the surface of the earth. How long would it take the group, traveling vertically down wards, to reach the earth and what fraction of them would remain undecayed by then?

Ans. 100 μ sec; practically zero.

31. In the laboratory, electrons from two accelerators are projected with the same speed of 2×10^8 m/sec but in opposite directions. What is the relative velocity of the two sets of electrons?
Ans. 2.77×10^8 m/sec.

32. A man on the moon observes two spaceships coming towards him from opposite directions at speeds of 0.8 c and 0.9 c respectively. What is the relative speed of the two space ships as measured by an observer on either one? **Ans. 0.987 c.**

33. Write a note on Einstein's mass-energy relation. What is the principle of mass and energy equivalence? Explain it by giving examples.

34. Show that the expression $E_K = \frac{1}{2}mv^2$ does not give the relativistic value of the Kinetic energy of a body even if m represents its relativistic mass.

35. Find the speed of a 0.1 MeV electron according to classical and relativistic mechanics.
Ans. 1.87×10^8 ms^{-1}; 1.64×10^8 ms^{-1}.

36. Compute the energy released on the conversion of 1 gm of matter into energy in KWh. (1 KWh = 3.6×10^6 J). **Ans. 2.5×10^7 KWh.**

37. A particle of rest mass m_0 is moving with a velocity of 0.9 c. Calculate (a) its relativistic mass, (b) its K.E.
Ans. (a) 2.3 m_0, (b) 1.3 m_0c^2.

38. A proton of rest mass 1.67×10^{-21} Kg is moving with velocity 0.9 c. Find its mass and momentum.
Ans. (a) 3.83×10^{-21} Kg. (b) 1.034×10^{-8} Kg ms^{-1}.

39. An electron is moving at a speed of 2.7×10^8 ms^{-1}. Find the ratio of its mass at this speed to its rest mass. Find also the total energy and the K.E. of the electron.
Ans. 2.29; 1.15 MeV; 0.65 MeV.

40. If a particle could move with the velocity of light, how much K.E. would it possess? **Ans. K.E. = ∞.**

41. (a) How much mass is lost by 1 Kg of water at 0°c when it turns to ice at 0°c?

(b) What change in mass is associated, in a chemical reaction, with (i) absorption (ii) release of 1 eV of energy.

Ans. (a) 3.7×10^{-12} Kg. (b) (i) increase of 0.18×10^{-32} gm. (ii) decrease of 0.18×10^{-32} gm.

CHAPTER XII

PHOTO-ELECTRIC EFFECT

Introduction

When radiations such as γ-rays, x-rays, ultraviolet rays and even visible light falls on a good number of substances, chiefly metals, electrons are ejected from these substances. This phenomenon is called *photo-electric effect.*

The phenomenon was first discovered by a telegraph operator, W. Smith, who using selenium resistors in the apparatus for the measurement of the resistance of transatlantic cables, observed that when sunlight fell upon the resistors the current in the circuit varied considerably. In 1887, Heinrich Hertz observed the same phenomenon accidently while working with resonance electric circuits in connection with electro-magnetic waves. He noted that when ultraviolet rays fell on a spark gap the sparks passed more easily. In 1888, Hallwachs, Elster and Geitel made the important observation that

(i) when ultra-violet light falls on a neutral zinc plate, the plate becomes positively charged.

(ii) when ultra-violet light falls on a negatively charged zinc plate, it loses its negative charge.

and (iii) when ultra-violet light falls on a positively charged zinc plate, it becomes more positively charged.

Hallwach *et. al.*, therefore came to the conclusion that only negatively charged particles can be emitted by the zinc plate when it is irradiated with ultra-violet light. [Afterwards, it was discovered that alkali metals like lithium, sodium, potassium, rubidium and caesium eject electrons even when ordinary light *i.e.*, visible light falls on them.]

P. Lenard and J.J. Thomson in 1900 measured the e/m (ratio of electric charge to mass) of the ejected ions and found them to be electrons. Although these electrons are no different from all other electrons, it is customary to refer to them as *photo-electrons*.

12.1 Experimental study of photo-electric effect

The phenomenon of photo-electric emission can be studied in detail with the apparatus shown in Fig. 12.1. In this figure, S is a source of radiation of variable and known frequency ν and intensity I. E is an emitting electrode of the material being studied and C is the

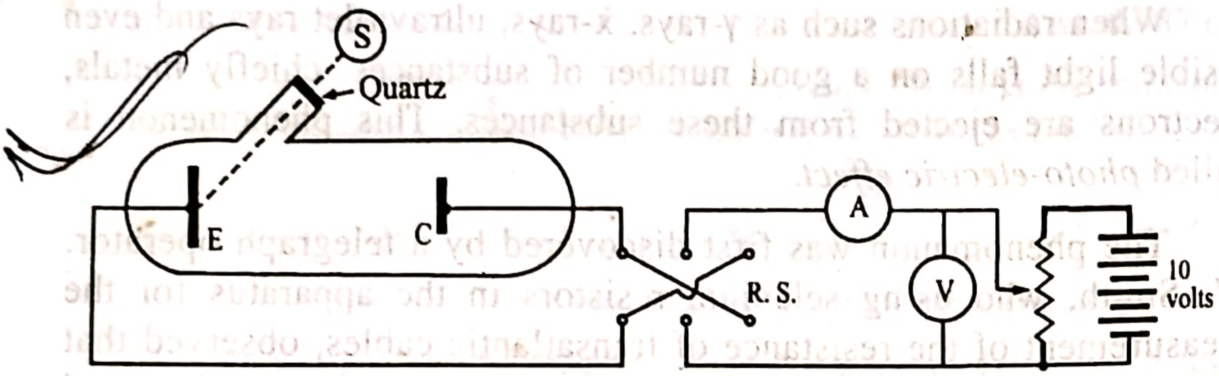


Fig. 12.1

collecting electrode. Both the electrodes are enclosed in an evacuated glass envelope with a quartz window that permits the passage of ultra-violet and visible light. As shown in the figure any potential difference can be established between E and C. A reversible switch R.S. enables the polarities of the two electrodes to be reversed. If the tube is in the dark, *i.e.*, no light falls on the electrode, there are no flow of current in the circuit and the microammeter reads zero. If ultra-violet light is allowed to fall on the emitting electrode, and if the collecting electrode is made positive with respect to the emitting electrode, current starts flowing as indicated by the microammeter. The explanation lies in the fact that when E is irradiated with ultra-violet light, electrons are liberated. As C is positive with respect to E, these photo-electrons will be quickly swept away from the emitter and attracted by the collector thereby starting the current flow. However, if the polarities between the plates are reversed with the help of the reversing switch R.S., no current flow is indicated by the ammeter. Although electrons are still being ejected by the emitter E, they cannot reach the plate C

(i) because of the pulling effect of the plate E which is now positively charged

and (ii) due to repulsion from the plate C which is now negatively charged.

This confirms the negative nature of the charge of the ejected particles.

Characteristics (laws) of photo-electric emission

From the experimental data collected by Richardson and Compton, the following fundamental laws regarding the emission of photo-electrons were established.

(i) By holding the frequency of radiation (ν) that falls on the photosensitive surface and the accelerating potential V constant, the intensity of the radiation is gradually increased and the corresponding photo-electric current is measured. When the photo-electric current is plotted against the corresponding intensity of radiation, a graph similar to one in Fig. 12.2 is obtained. As can be seen

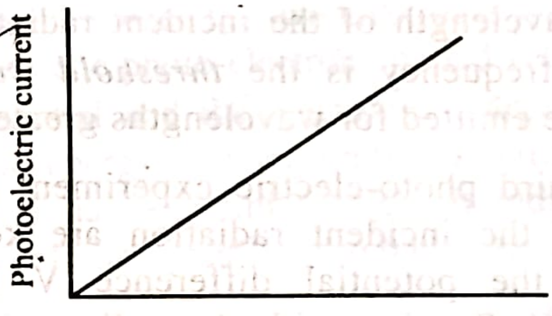


Fig. 12.2

from the graph, *the strength of the photo-electric* current (*i.e.*, the number of electrons emitted per second) *is directly proportional to the intensity of light or radiation used, provided the frequency of the radiation is kept constant.*

(ii) Next the intensity of light or radiation was kept constant. Starting from zero, the frequency of the radiation was gradually increased. At first there was no photo-electric emission and hence no current was indicated in the ammeter. This continued until a certain value of the frequency was reached when photo-electrons began to be emitted, registering a current in the ammeter. With further increase in frequency, the photo-electric current is found to increase although non-linearly. When the result is represented graphically, curves similar to those shown in Fig. 12.3 were obtained. A and B in Fig. 12.3, represent two different irradiated materials. The significant thing about these curves is that for every substance

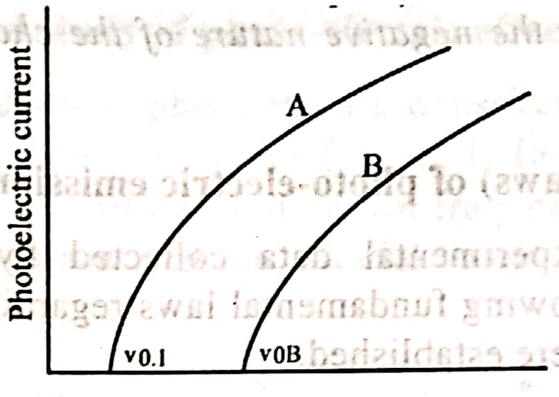


Fig. 12.3

irradiated there is a limiting or critical frequency below which no photo-electrons are emitted. This limiting frequency is called the *threshold frequency*, ν_o, and is a characteristic of the material irradiated *i.e.*, its value depends on the nature of the material irradiated. The wavelength of the incident radiation corresponding to the threshold frequency is the *threshold wavelength*, λ_o. No photo-electrons are emitted for wavelengths greater than this.

(iii) In the third photo-electric experiment both the intensity and frequency of the incident radiation are kept constant. The variable now is the potential difference V between the two electrodes E and C. Starting with the collector at about 10 volts positive, the potential is gradually reduced to zero and then run negative until the photo-electric current stops entirely. Curve I_1 in Fig. 12.4 shows the type of curve expected for a particular substance. The curve requires careful interpretation.

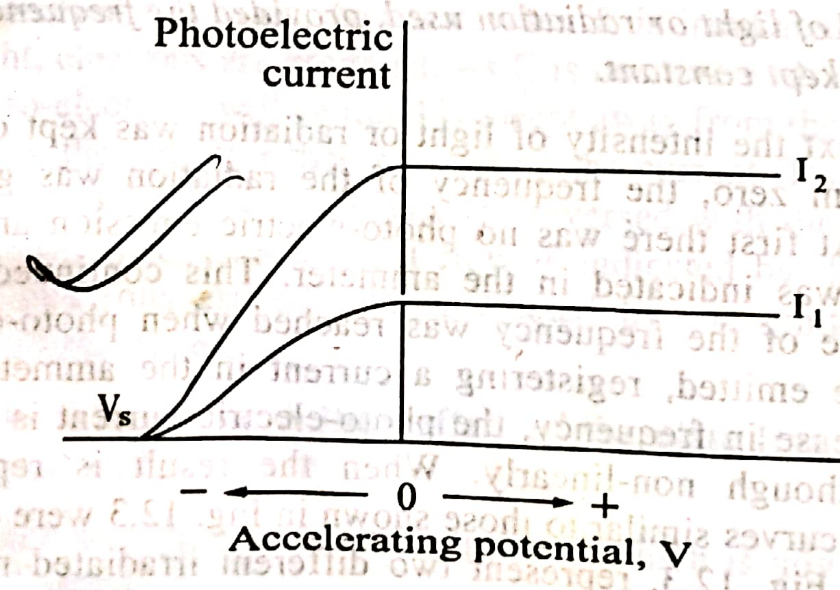


Fig. 12.4

When the potential difference between E and C is about 10 V or more, all the emitted electrons are immediately collected by C. This is the limiting or *saturation* value of the current. Further increase in V hardly produces any appreciable increase in current as shown by the flat portion of the curve. As the accelerating potential is now reduced from positive values through zero to negative values, the photo-electric current does not immediately drop to zero. This proves that electrons are emitted from E with some definite velocity-sufficient enough to give kinetic energy to the electrons so as to surmount (overcome) the retarding (opposing) electric field between the electrodes. Hence, some electrons still manage to reach C despite the fact that the electric field oppose their motion.

Eventually when the negative potential is made large enough, a value V_S is reached when the current is reduced to zero. V_S is known as the *stopping potential* and is defined as that value of the retarding potential difference between the two electrodes which is just sufficient to halt the most energetic photo-electron emitted. Therefore, the product of the stopping potential and the electronic charge, $V_S.e$, is equal to the maximum kinetic energy that an emitted electron can have. Since the stopping potential has a definite value, it indicates that the emitted electrons have a definite upper limit to their kinetic energy.

Doubling the intensity of light merely doubles the current at each potential as shown by curve I_2, but does not affect the value of V_S. Or, *the stopping potential, V_S, is independent of the intensity of incident radiation.*

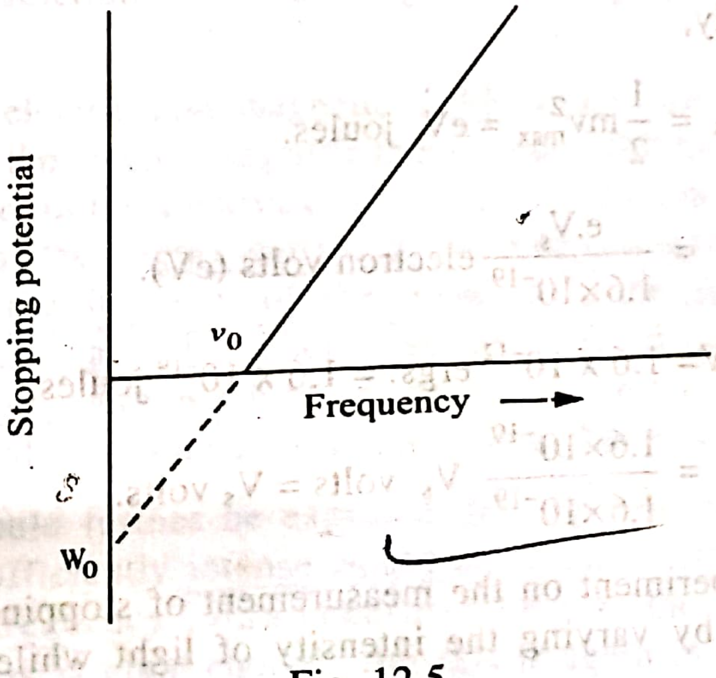


Fig. 12.5

If, however, the experiment is repeated by varying the frequency of light, it is found that the stopping potential varies linearly with the frequency as shown in Fig. 12.5. As no electrons are emitted below the threshold frequency, the stopping potential is zero for that region. But as the frequency is increased above the threshold frequency, *the stopping potential varies linearly with the frequency of the incident light.*

The maximum velocity, v_{max}, of emission of a photo-electron can be determined from the value of the stopping potential, V_s, as follows:

$$\frac{1}{2}mv_{max}^2 = e.V_s$$

where m and e are the mass and charge of an electron.

$$\text{or,} \quad v_{max} = \sqrt{\frac{2eV_s}{m}}$$

$$= \sqrt{\frac{2\times1.602\times10^{-19}\times V_s}{9.109\times10^{-31}}}$$

$$= \sqrt{2\times1.602\times10^{11}\times V_s}$$

$$= 5.93\times10^5\sqrt{V_s} \text{ metre/second.}$$

Obviously,

$$E_{max} = \frac{1}{2}mv_{max}^2 = eV_s \text{ joules.}$$

$$= \frac{e.V_s}{1.6\times10^{-19}} \text{ electron volts (eV).}$$

since 1 eV= 1.6×10^{-12} ergs. = 1.6×10^{-19} joules.

$$\therefore \quad E_{max} = \frac{1.6\times10^{-19}}{1.6\times10^{-19}} V_s \text{ volts} = V_s \text{ volts.}$$

If the experiment on the measurement of stopping potential is repeated first by varying the intensity of light while keeping the

frequency of emission constant and then by varying the frequency while keeping the intensity constant, the following important conclusion emerges.

"The maximum velocity of an electron emitted (and hence its kinetic energy) varies linearly with the frequency of the incident light but is independent of its intensity."

From the measurements for stopping potential it was further established that *the velocities of emitted electrons have values between zero and a definite maximum.* The proportion of the electrons having a particular velocity is independent of the intensity of radiation.

(iv) It was also established experimentally that the *photo-electric emission is an instantaneous process.* If there is any time lag between the arrival of light at a metal surface and the emission of photo-electrons, it should be less than 3×10^{-9} seconds which is the limit of experimental accuracy.

Failure of classical electro-magnetic theory to explain photo-electric effect

Although at the first glance the electro-magnetic theory of light seems to be capable of explaining the photo-electric effect, it soon becomes obvious that the main features of photo-electric emission leads to contradictions if electro-magnetic theory is applied to the problem.

(i) The electric and magnetic fields which are attributed to light waves in the electro-magnetic theory, may well be able to exert sufficient force on the electrons in the metal and release them from its surface. But the energy carried by an electro-magnetic wave is described by the intensity of the wave, *i.e.*, the square of the amplitude of the wave. The photo-electrons ejected from the metal should then have energies which are dependent on the intensity, but the experimental results described above show that this is not so.

(ii) It would further be expected that light of low frequency (infrared) if sufficiently intense would be as much effective as the light of high frequency (ultra-violet), as long as the intensity and time of exposure is great enough. But this is again in contradiction

to the experimental results which show that light of frequency below a sharply defined threshold is absolutely incapable of producing any effect.

(iii) Equally difficult to explain was the fact that there is no time lag between the arrival of the radiation on the metal surface and the emission of photo-electrons. For example let us consider an experiment as depicted in Fig. 12.1 where violet light falls on a sodium surface. A detectable photo-electric current will result when 10^{-6} watt/m^2 of electromagnetic energy is absorbed by the surface. Since sodium is a good reflector of light, a more intense beam will be required in practice. Now there are about 10^{19} atoms in a layer of sodium one atom thick and 1 m^2 in area. If it is assumed that the incident light is absorbed in the 10 uppermost layers of sodium atoms, the 10^{-6} watt/m^2 is distributed among 10^{20} atoms. Hence each atom receives energy at the average rate of 10^{-26} watt, which is less than 10^{-7} eV/sec. It should therefore take more than 10^7 seconds, or almost a year, for any single electron to accumulate the 1 eV or so of energy that the photo-electrons are found to possess. In the maximum possible time of 3×10^{-9} second, an average electron according to electro-magnetic theory, will have gained only 3×10^{-16} eV. Even if some kind of resonance process is called upon to explain why some electrons acquire more energy than others, the fortunate electrons will still have no more than 10^{-10} of the observed energy.

12.2 The quantum theory of light

Thus the electromagnetic theory which could so successfully explain other phenomena exhibited by light such as interference, diffraction and polarization, completely fails to explain the photo-electric effect. In 1905, Albert Einstein found that the paradox presented by the photo-electric effect could be understood only by taking seriously a notion proposed five years earlier by the German theoretical physicist Max Planck. Planck was seeking to explain the characteristics of radiations emitted by black bodies – bodies hot enough to be luminous. The different formulae for energy distribution in black body radiation, derived on the basis of classical theory, not only led to wrong and even absurd conclusions, but also could not stand the test of observation. Planck realized that the failure of the above formulae was due to the assumption that the energy changes of radiators take place continuously. He therefore

proposed a new theory – the *quantum theory* according to which the radiation of energy from a system or the exchange of radiant energy between different systems occurs not in continuous fashion permitting all possible values as demanded by the wave theory but *discontinuously* as little bursts (or bundles) of energy, which are integral multiples of an elementary quantum of energy. These bursts of energy are called quanta. Planck found that the quanta associated with a particular frequency of light all have the same energy and that this energy E is directly proportional to the frequency ν of the incident radiation. That is

$$E = h\nu$$

where h is a constant, known today as Planck's constant, and has the value

$$h = 6.63 \times 10^{-34} \text{ joule-sec.}$$

[**Note:** Since the frequency ν is different for different radiations, the quantum unit hν is not the same for all kinds of radiation. In the original quantum theory the energy of the photon was always considered to be an integral multiple of the unit hν, as hν, 2hν, 3hν, ... But in the new quantum mechanics it is further refined to a value of $\left(n+\frac{1}{2}\right)h\nu$, n being an integer, which means that the limiting value of the energy of the photon is not zero, but $\frac{1}{2}h\nu$]

12.3 Einstein's photo-electric equation

While Planck had to assume that a hot body radiates energy intermittently in the form of a bundle of energy (quanta), he did not doubt that it propagates continuously through space as electromagnetic waves. According to wave theory of light pulsating electromagnetic fields spread out from their source. Einstein proposed that light not only is emitted a quantum at a time, but also propagates as individual quanta – radiation is not only emitted or absorbed in discrete amounts of quanta, but also the same quantum structure is retained by radiation while travelling through space, very much like a shot fired from a gun, without spreading out as waves in any supposed ether.

Einstein applied this theory to the photo-electric process and obtained a consistent and satisfactory explanation of all the experimental facts of photo-electric emission. In his theory, radiation is regarded as a shower of photons each of energy $h\nu$ moving in space with the velocity of light. When a single photon is incident on a metal surface, it is completely absorbed by an atom. The energy is subsequently imparted to one of the electrons of the atom. This energy is utilized for two purposes:

(i) partly for getting the electron free from the atom and away from the metal surface. This energy is known as the *photo-electric work function* of the metal and is represented by either φ or w_o.

(ii) the balance of the photon energy is used up in imparting to the freed electron a kinetic energy of $\frac{1}{2}mv^2$.

Einstein expressed this assumption in the form

$$h\nu = w_o + \frac{1}{2}mv^2 \qquad (12.1)$$

where

$h\nu$ = energy content of each quantum of the incident light.

w_o = photo-electric work function of the metal surface being irradiated.

$\frac{1}{2}mv^2$ = kinetic energy of the ejected photo-electron.

Expression 12.1 is known as *Einstein's photo-electric equation.*

Explanation of the characteristics of photo-electric equation

Einstein's photo-electric equation shows clearly why

(i) *the maximum velocity of the ejected photo-electron is directly proportional to the frequency of the incident radiation.*

Eqn. 12.1 can be written as

$$\frac{1}{2}mv^2 = h\nu - w_o$$

Since w_o is constant for a given emitter, $\frac{1}{2}mv^2$ and hence v, the velocity of the ejected electron, increases as ν increases.

(ii) *The velocity of the photo-electron is independent of the intensity of the radiation.*

If ν remains constant, increasing intensity merely increases the number of photons striking the emitter. This results in greater number of collisions resulting in the emission of greater number of photo-electrons. But each photo-electron will receive exactly the same energy ($h\nu - w_o$) and hence would acquire the same velocity irrespective of the intensity.

However, increase in the number of photo-electrons emitted results in the increase in the strength of photo-electric current. This explains why photo-electric current varies directly as the intensity of radiation.

(iii) *Existence of a threshold frequency that varies with the nature of the emitter.*

An electron is electro-statically bound to the emitter. The amount of energy required to free this electron from the emitter is known as the photo-electric work function, w_o, and this comes from the energy of the incident radiation $h\nu$. As ν is reduced the energy of the incident radiation is reduced also. At a value, say ν_o, let $h\nu_o = w_o$. The photon energy is then just sufficient to liberate the electron only and no energy will be available for imparting kinetic energy to the electron. Hence eqn. 12.1 would reduce to

$$h\nu_o = w_o$$

ν_o is called the *threshold frequency* and is defined *as the minimum frequency which can cause photo-electric emission.* For frequencies lower than ν_o the energy of the incident photon is less than the work function of the material and hence there would be no emission of electrons. But for frequencies greater than ν_o, electrons would be emitted with a certain definite velocity (and hence kinetic energy). This explains why there is a threshold frequency ν_o connected with photo-electric emission.

(iv) *There is no time-lag in the process of photo-electric emission.*

As the phenomenon of photo-electric emission is considered as a collision process between two particles – the photon and the atom, there is no time-lag between the incidence of the photons on the metal surface and the emission of electrons therefrom.

(v) *Emission of photo-electrons with all possible velocities upto a certain maximum.*

w_o is the work that must be done to take an electron through the metal surface from just beneath it; more work is required when the electron originates deeper in the metal. Thus electrons from various depths below the metal surface would emerge with different energies *i.e.*, velocities. The greater the depth the less is the velocity. It is therefore easy to see why not all the photo-electrons have the same energy, but emerge with all possible velocities upto a certain maximum. For an electron just beneath the metal surface the kinetic energy will be maximum. Eqn. 12.1 then becomes

$$h\nu = w_o + \frac{1}{2}mv^2_{max}$$

$$= w_o + (K.E.)_{max}$$

Long Wavelength Limit (λ_o)

The wavelength corresponding to the threshold frequency ν_o is referred to as the long wavelength limit λ_o. Radiations having wavelength longer than λ_o would not be able to eject electrons from a given material whereas those having $\lambda < \lambda_o$ will. In other words, it represents the upper limit of wavelength for photo-electric emission. By analogy, it is also referred to as *threshold wavelength.*

Now $c = \nu_o\lambda_o$ or, $\lambda_o = \dfrac{c}{\nu_o}$

Also $w_o = h\nu_o$ or, $\dfrac{1}{\nu_o} = \dfrac{h}{w_o}$

$$\therefore \quad \lambda_o = \frac{c}{\nu_o} = \frac{ch}{w_o}$$

(i) when w_o is in joules

$$\lambda_o = \frac{3\times10^8\times6.625\times10^{-34}}{w_o}$$

$$= \frac{19.875\times10^{-26}}{w_o} \text{ metre}$$

(ii) when w_o is in electron-volts (eV)

$$\lambda_o = \frac{3\times10^8\times6.625\times10^{-34}}{1.602\times10^{-19}\times w_o}$$

$$= \frac{12.4\times10^{-7}}{w_o} \text{ metre}$$

$$= \frac{12,400}{w_o} \text{ Å}.$$

Kinetic energy of photo-electrons

From Einstein's photo-electric equation we get

$$h\nu = w_o + \frac{1}{2}mv^2$$

or, $\frac{1}{2}mv^2 = h\nu - w_o.$

Now w_o is the energy necessary to free the electron from the atom. In case ν_o is the threshold frequency for a particular metal, then all the energy of the incident photon is just sufficient to liberate the electron only and no energy is available to impart kinetic energy to the electron.

In that case

$$w_o = h\nu_o$$

for the particular metal

Thus $\frac{1}{2}mv^2 = h\nu - h\nu_o$

or, K. E. = $h\nu - h\nu_o$.

$= h(\nu - \nu_o)$

$= ch\left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right)$ since $\nu = \frac{c}{\lambda}$

$$= 3\times10^{8}\times6.625\times10^{-34}\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$$

where $c = 3 \times 10^8$ m/sec

and h = Planck's constant

$= 6.625 \times 10^{-34}$ joules-sec.

$\therefore$ K. E. $= 19.875 \times 10^{-26}\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$ joules ; λ and λ_o in metres.

$= \frac{19.875\times10^{-26}}{10^{-10}}\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$ joules ; λ and λ_o in Angstrom unit (1 A.U. = 10^{-10}m)

$= 19.875 \times 10^{-16}\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$ joules ; λ and λ_o in A.U.

$= \frac{19.875\times10^{-16}}{1.602\times10^{-19}}\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$ eV ; λ and λ_o in A.U.

$$\therefore \text{K. E.} = 12{,}400\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)\text{eV} \quad (12.2)$$

where λ and λ_o are expressed in Angstrom unit ($\mathring{A}$).

The kinetic energy referred to in eqn. 12.2 above represents the maximum kinetic energy that an electron may possess.

E_{max} can also be obtained from the expression given below.

$E_{max} = h\nu - h\nu_o = h(\nu - \nu_o) = h.\Delta\nu$ joules.

Photo-electric work function

As already explained, the photo-electric work function is defined as the energy which is just sufficient to liberate electrons from a body with zero velocity. Its value is given by

$$w_o = h\nu_o = \frac{ch}{\lambda_o}$$

$$= \frac{3\times10^8 \times 6.625\times10^{-34}}{\lambda_o} \text{ joules}$$

$$= \frac{19.875\times10^{-26}}{\lambda_o} \text{ joules ; } \lambda_o \text{ in metres}$$

$$= \frac{19.875\times10^{-16}}{\lambda_o} \text{ joules ; } \lambda_o \text{ in } \mathring{A}$$

$$= \frac{19.875\times10^{-16}}{1.602\times10^{-19}\lambda_o} \text{ eV ; } \lambda_o \text{ in } \mathring{A}$$

$$\therefore \quad w_o = \frac{12,400}{\lambda_o} \text{ eV ; } \lambda_o \text{ in } \mathring{A}. \qquad (12.3)$$

12.4 Millikan's experiment

To verify Einstein's photo-electric equation and to determine the value of the Planck's constant h and the photo-electric work function w_o Millikan performed an experiment in 1916. The experiment is based on the determination of stopping potential. Millikan's experimental set up is shown in Fig. 12.6. The apparatus consists of an evacuated chamber C. At the centre of the chamber is kept a drum D which can rotate freely about a vertical axis. Four

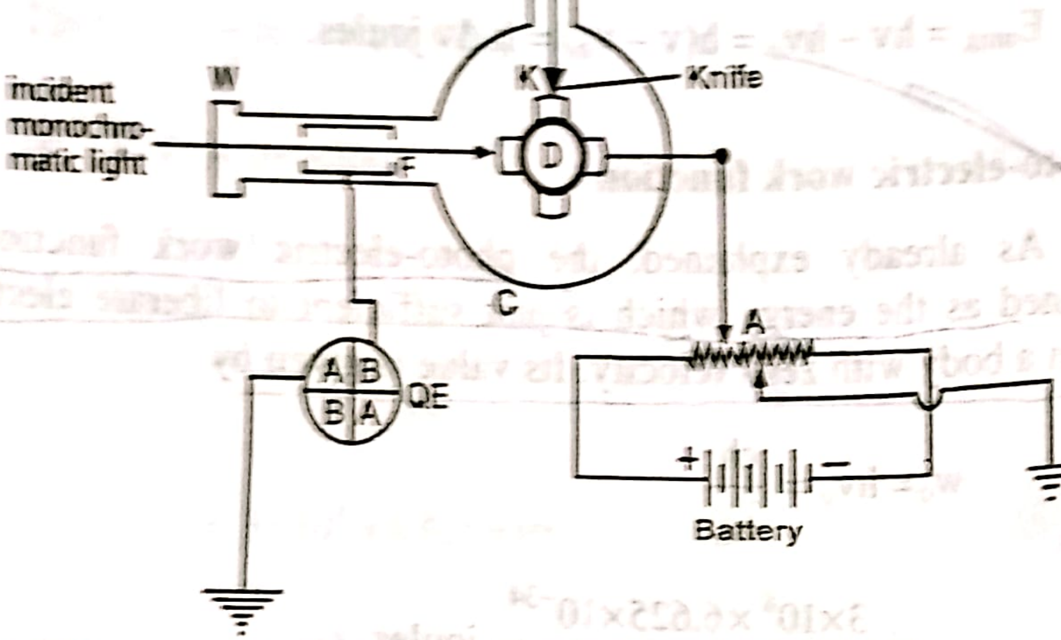


Fig. 12.6

cylindrical blocks of alkali metals like sodium, potassium, lithium, etc. are fixed on the periphery of the drum. Alkali metals were preferred as they readily exhibit photo-electric emission even with visible light. By rotating the drum, any one of the metal blocks could be turned towards a quartz window W through which monochromatic light of known frequency is allowed to pass. As these metals gets oxidized easily, the experiment was carried out in vacuum. Further, fresh surfaces of the metals were obtained by scraping the metals with a sharp knife-edge K which could be brought opposite the metal block by means of an electro-magnet placed outside the chamber. F is a Faraday cylinder connected to a quadrant electrometer QE. With the help of a sensitive potential divider arrangement suitable positive or negative potential can be applied to the drum with respect to the cylinder F.

To start with, the surface of any one of the metal blocks is scraped with the knife-edge K and then by rotating the drum, the block is turned towards the window W. Monochromatic radiation of known frequency ν is then allowed to be incident on the surface of the metal block. If the drum is kept at a negative potential with respect to the cylinder F, the photo-electrons ejected from the surface are accelerated towards the cylinder thereby recording a deflection in the quadrant electrometer. When a small positive potential is applied to the drum to prevent the

photo-electrons from leaving the drum, only those fast moving photo-electrons *i.e.*, photo-electrons having sufficient kinetic energy to overcome this opposing retarding potential reach the cylinder F. The deflection in the electrometer therefore decreases. The positive potential of the drum is gradually increased till the deflection in the electrometer becomes zero. At this stage no photo-electron can reach the cylinder. This particular potential V_s for which no electron can reach the cylinder is called the stopping potential. The stopping potential under these conditions is the positive potential applied to the drum which corresponds to zero current in the electrometer.

Even when the intensity of light was increased by keeping the frequency constant, the deflection becomes zero at the same potential. *This means that at a constant frequency of incident radiation* the stopping potential V_s is the same and is independent of the intensity of the incident radiation. The experiment was repeated with different lights *i.e.*, radiation with different frequencies. It was found that the stopping potential V_s increases with increase in frequency for the same metal. According to Einstein's photo-electric equation

$$h\nu - w_o = \frac{1}{2}mv^2$$

$$\text{or,} \quad h\nu - h\nu_o = \frac{1}{2}mv^2 \qquad [w_o = h\nu_o]$$

$$\text{Also } eV_s = \frac{1}{2}mv^2$$

$$\therefore \quad eV_s = h\nu - h\nu_o$$

$$= h(\nu - \nu_o)$$

$$\therefore \quad V_s = \frac{h}{e}(\nu - \nu_o)$$

The experiment was repeated for different metals and graphs were drawn with ν along the X-axis and V_s along the Y-axis. The resulting graph was a straight line in all cases and this is in agreement with Einstein's theory.

Determination of Plank's constant

The stopping potentials for a particular metal were determined for different frequencies in the manner mentioned above. A graph is then plotted with ν along the X-axis and V_s along the Y-axis. The resulting graph was a straight line as shown in Fig. 12.7(a). Similar graphs were obtained for cesium, and zinc as shown in Fig. 12.7(b). The point at which the straight line AB intercepts the X-axis gives the threshold or cut off or critical frequency ν_0, and varies from

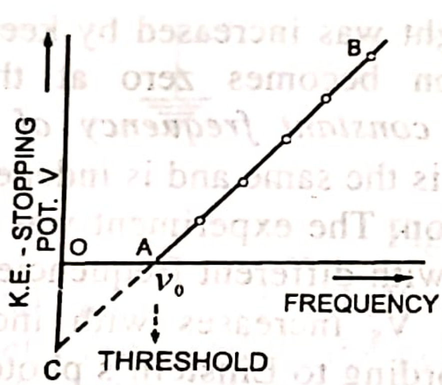


Fig. 12.7(a)

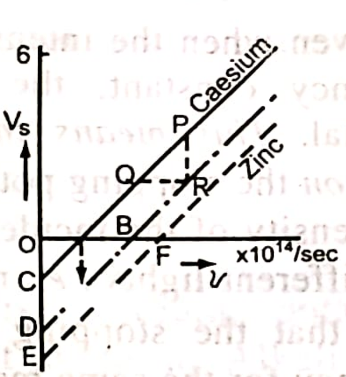


Fig.12.7(b)

metal to metal as B and F in Fig. 12.7(b). For this frequency the kinetic energy of the emitted photoelectron is zero ($\frac{1}{2}mv^2 = 0$) and the energy of the photon $h\nu_0 = w_0$. At a frequency less than ν_0 no electrons are emitted from the surface. The slope of the graph

$$= \frac{PR}{QR} = \frac{V_s}{\nu - \nu_0}$$

But the equation of the graph may be expressed as

$$e.V_s = h\nu - h\nu_0 \qquad (12.4)$$

$$\text{or,} \quad \frac{1}{2}mv^2 = h\nu - h\nu_0$$

from eqn. 12.4

$$\frac{V_s}{\nu - \nu_0} = \frac{h}{e}$$

Therefore, the slope $\frac{PR}{QR} = \frac{h}{e}$ on the experimental graph; then $\frac{h}{e} = \tan\theta$. Thus knowing the value of e, h can be calculated from the relation

$$h = e. [\text{ slope of the graph }]$$

By experimental determination of the stopping potential for a metal, and the frequency of the incident light and also the threshold frequency, it was found that the value of h turns out to be 6.62×10^{-34} J-sec when $e = 1.6 \times 10^{-19}$ coulomb. This value is in agreement with the value of the Planck's constant as determined by other methods. Millikan's experiment is therefore a confirmation of Einstein's photo-electric equation and hence quantum theory of radiation on which it is based. The intercept OC of the line with the vertical axis gives the value of $\frac{w_o}{e}$ where $w_o = h\nu_o$ is the work function of the metal.

Hence

$$w_o = e \times OC$$

Experimentally it is found that the work function w_o for barium = 2.5 eV, for sodium = 2.46 eV, for potassium = 2.24 eV and for cesium = 1.92 eV.

From what has been discussed above it follows that the graphs in Fig. 12.7 can be used to determine the value of (i) the Planck's constant h, (ii) threshold frequency ν_o of a metal and (iii) the work function w_o for the metal.

12.5 Photo-electric cell

Photoelectric cell is an arrangement to convert light energy into electrical energy. The following three main types of photo-electric cells are worth considering :

(i) Photoemissive cell : it depends on the mission of electrons from a metal cathode when it is exposed to light or other radiation.

(ii) Photovoltaic cell : the sensitive element here is a semiconductor – not a metal. It generates voltage in proportion to the light or any radiant energy incident on it.

(iii) Photoconductive cell : it depends on the change of resistance of a semiconductor material in accordance with the radiant energy received.

Photo-emissive cell or Phototube:

construction

it consists of two *metallic* electrodes – a cathode and an anode,

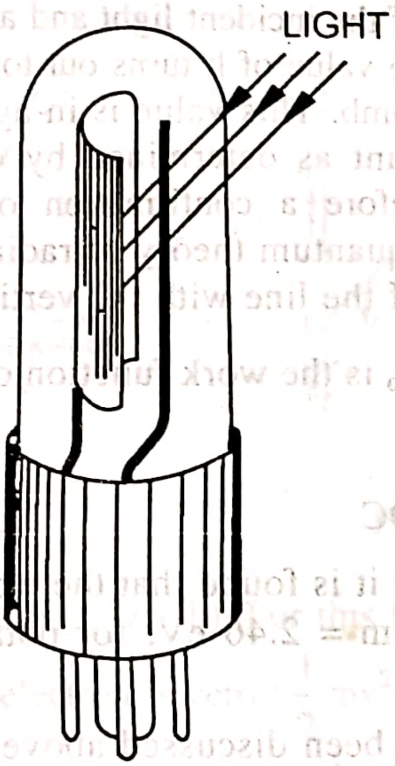


Fig. 12.8

enclosed in an evacuated glass bulb fitted with a base like a thermionic value as shown in Fig. 12.8. The cathode is either V-shaped or semi-cylindrical metal plate coated with an emissive material and is known as the emitter. The anode is in the form of a thin wire of platinum or nickel fixed along the axis of the cylinder or it can be a small button so placed as not to obscure the light source from the cathode. This wire is known as the collector.

Various types of photocathodes in use are

(i) pure metals like Ni, Zn, Al etc. for ultra-violet radiations.

(ii) alkali metals for visible light.

(iii) composite cathodes for infra-red radiations such as a silver plate coated with caesium or potassium.

(iv) alloy cathodes made up of thin films of antimony caesium.

working

when radiation falls on the cathode, electrons are emitted and are attracted by the positive anode. Hence, a current is produced whose magnitude, for a given cathode, depends on

(i) intensity of incident radiation

and (ii) anode-to-cathode voltage.

The variations are shown in Fig. 12.9. As can be seen, the current produced by such a cell is of the order of a few microamperes. The

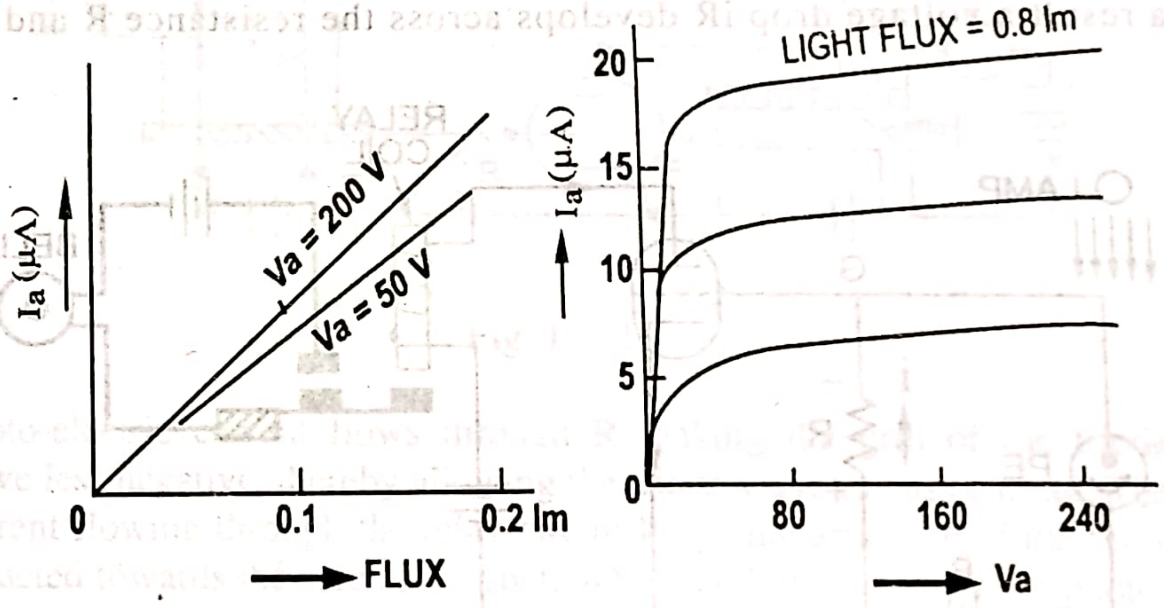


Fig. 12.9

current, therefore, has to be further amplified before being put to use.

As the current in a vacuum type photo-cell is very weak, gas filled type of photo-cells are used. The gas used is either argon or helium. When bombarded by photo-electrons, the gas molecules are ionized; resulting in a larger current flow in the circuit.

uses

photo-electric cells find wide use in the field of photometry – for accurate comparison of light intensities, in calorimetry, television, burglar alarms, fire alarms, counting machines, complexion metres and traffic regulators. They are used for colour identification, for reproduction of sound from a motion-picture film and for automatic switching on and off of street lights. They are also used for *on and off* circuits and in other circuits concerning the counting or sorting of objects on a conveyer belt, the automatic opening of a door as it is approached, etc.

Automatic Bell Alarm Circuit

Fig. 12.10 shows the circuit of an automatic bell alarm circuit in which a bell rings up whenever light beam between the lamp and the phototube is interrupted. Whenever light falls on the phototube, it causes a small current to flow from battery through resistance R. As a result a voltage drop iR develops across the resistance R and is

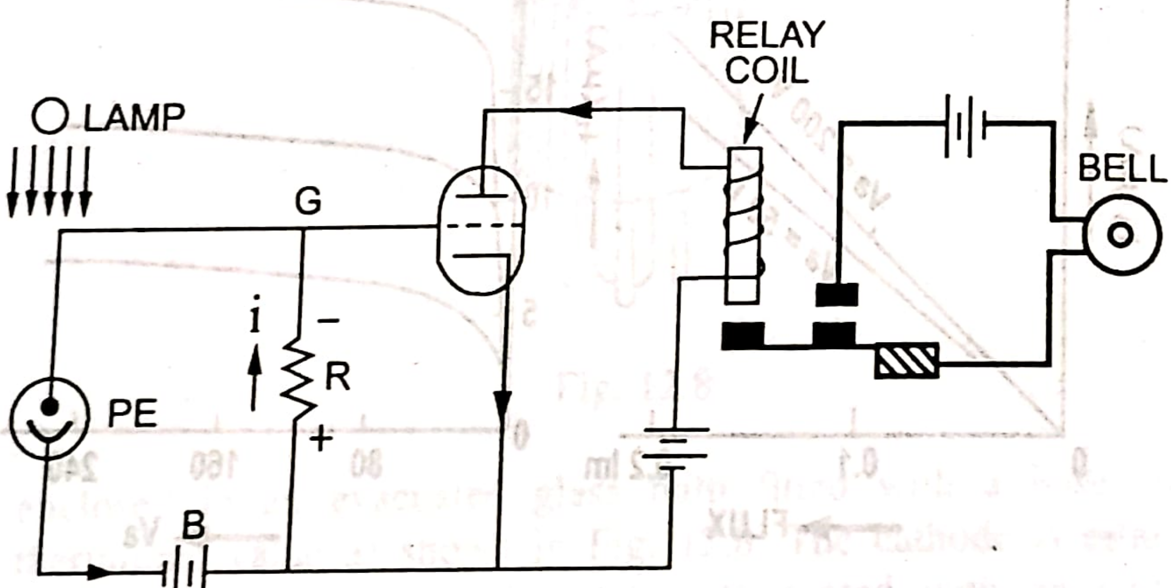


Fig. 12.10

applied to the grid G of a triode value. The grid G is maintained negative with respect to the cathode. Due to this negative bias to the grid, very little or no anode current flows. Should the light weaken or be cut off, the grid bias reduces and allows more current to flow. The anode current flows through the coil of a sensitive relay. Because of this current flow, the armature of the relay is attracted

thus closing the contacts of a circuit-consisting of a battery and an electric bell or buzzer. If the photocathode is made sensitive to the infra-red or ultra-violet light, then there is no need to use visible light.

Photo-electric relay

One of the most important uses of the photoelectric cell is to operate a relay which automatically opens or closes a local circuit. This can be employed for a variety of useful purposes.

Fig. 12.11 shows the circuit. With no light incident on the phototube, the grid G, of the triode is maintained sufficiently negative by the battery to cut off the plate current. However, when light is incident on the photo-electric cell, photoelectrons are ejected out of the cylinder B. The resulting

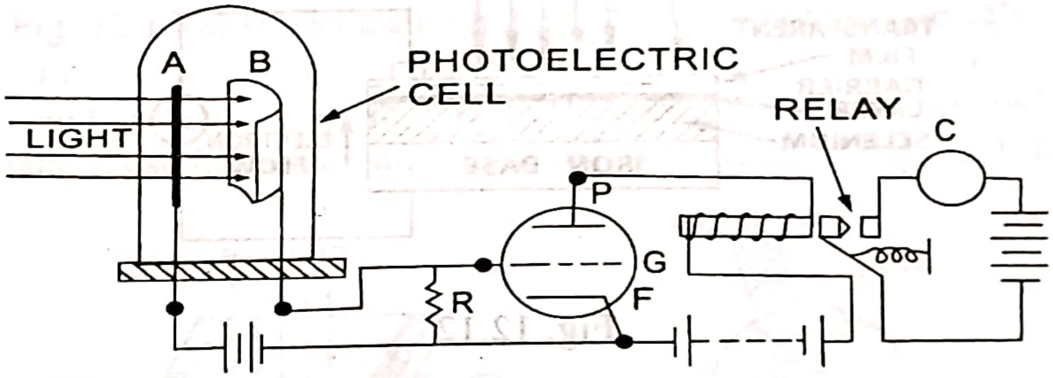


Fig. 12.11

photo-electric current flows through R making the grid of the triode valve less negative, thereby allowing the triode valve to pass current. The current flowing through the relay circuit keeps the armature of the relay attracted towards the electro-magnet. When the light is cut off, no photo-electrons are ejected and hence no photo-electric current. The grid becomes more negative in the absence of the photo-electric current and hence no current is allowed to flow in the plate circuit. The armature is no longer attracted by the electro-magnet and it springs back to make contact with D. The local circuit closes and the current flows through it. If C is a bell it starts ringing and if C is a bulb, it glows. It is therefore obvious that so long as the photo tube is illuminated, the relay remains inoperative but it starts operating immediately on interruption of the light. A number of industrial operations are performed in this way.

Photo-voltaic cell

In the case of photo-electric cells, a positive potential is to be applied to the collector to attract the electrons and therefore an external battery is a necessity. But the photo-voltaic cell is a self-generating cell where the electrons ejected by light, themselves produce a potential difference between the two plates which causes the current to flow through the external circuit even though no external battery is applied.

The most commonly used photo voltaic cells are of the barrier layer type like iron-selenium cells or $Cu - CuO_2$ cells. In the iron-selenium cell, selenium layer is placed on an iron disc (Fig. 12.12).

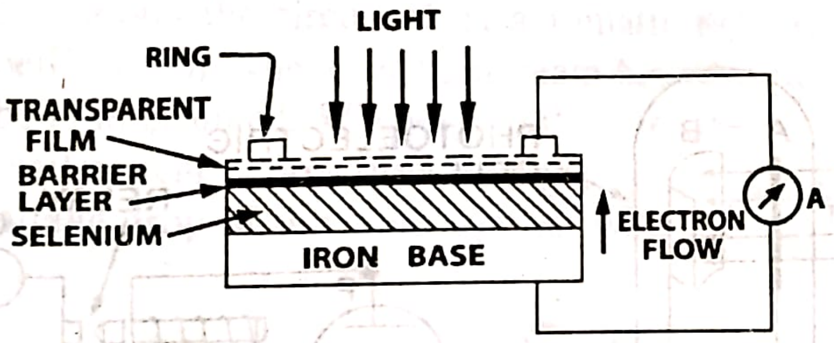


Fig. 12.12

An extremely thin semi-transparent layer of gold or silver is formed on the selenium to act as a front electrode. The barrier layer is formed by cathode-sputtering the semi-transparent film on the selenium. A contact ring on the silver layer acts as one electrode and the iron base as the other.

When light or radiation falls on the semiconductor *i.e.*, selenium, it ejects electrons which travel from selenium to the front silver electrode as shown in Fig. 12.12. As the (boundary) barrier layer acts as a rectifier, it does not permit the flow of electrons in the opposite direction. The e.m.f. generated internally between silver electrode and selenium is almost directly proportional to the incident flux.

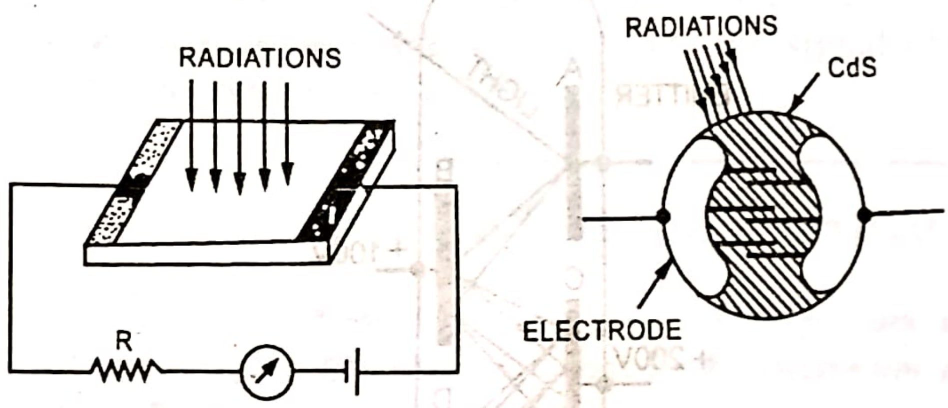
The main advantage of a photo-voltaic cell is that *no external battery is required for its operation i.e.,* it is self-generating. Moreover, the internal e.m.f. and hence current generated by it are

large enough to be measured on a pointer galvanometer. Hence such cells can be calibrated and used in devices like *light meters* in photography and direct-reading illumination metres. With low resistance relays, photo-voltaic cells can also be used for on/off operations and other monitoring operations in industry.

Photoconductive cell

Photoconductive cell is based on the principle that resistivity of semiconductor materials like selenium (Se), cadmium sulphide (CdS), lead sulphide (PbS) and thallium sulphide (TlS) decreases when irradiated. In other words, such materials have high *dark* resistance and low *irradiated* resistance. In some cases the resistance drops from dark values of 10 to 20 million ohms to approximately one million ohms when exposed to light.

Fig. 12.13(a) shows a simple form of photo-conductive cell using selenium. It consists of the semi-conductor material with two electrodes attached to it. When the cell is unilluminated, its resistance is so high that current through the circuit of Fig. 12.13(a) is



(a) (b)

Fig. 12.13

very low. When radiation is allowed to fall on the cell, its resistance decreases and the current through the circuit becomes large. The shape of the semiconductor material is so made as to obtain a large ratio of *dark to light* resistance.

Fig. 12.13(b) shows a commonly used CdS cell. It has a very high *dark to light* ratio and gives maximum response at 5000 A.U. As shown in the figure, the two electrodes are extended in an inter digital pattern in order to increase the contact area with the sensitive material.

Photoconductive cells using TlS and PbS have been used for detection of ships and aircrafts by the radiations given out by their exhausts or tunnels. They have also been used for telephony by modulated infra-red light.

Photo-multiplier tube

Photomultiplier tube is a device in which current is amplified by electron multiplication through secondary emission. When a high energy electron strikes the surface of a metal plate, preferably an alkali metal, it knocks free additional electrons from the surface. This process is known as *secondary emission* and the electrons thus emitted are known as *secondary electrons*. In case of cesium and potassium, one primary electron can knock out about eight to ten electrons from the surface.

When light falls on the cathode or the emitter A (Fig. 12.14), photo-electrons are ejected from it. Instead of being collected by the

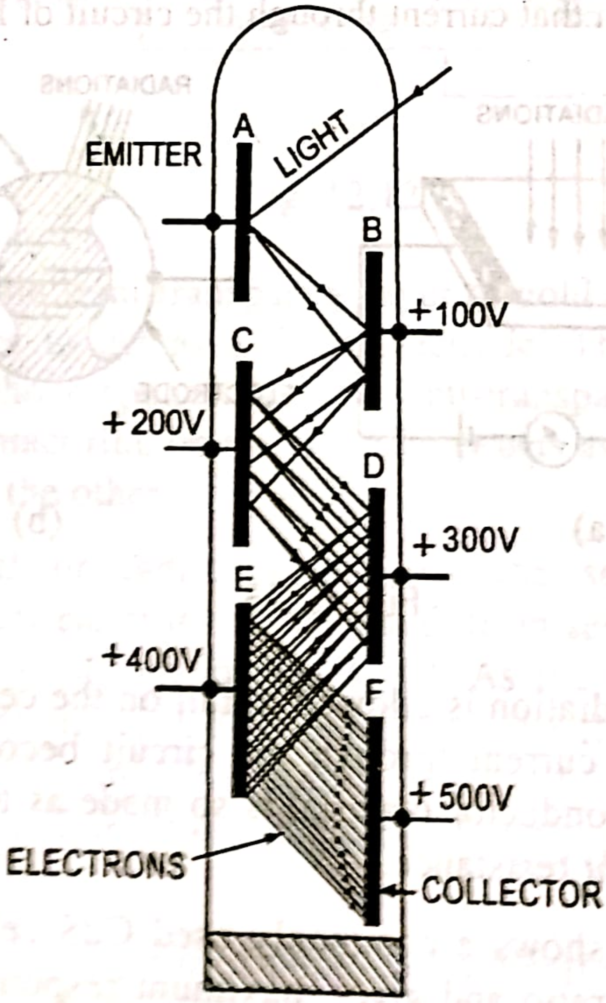


Fig. 12.14

anode, these electrons are directed towards a secondary electron emitting electrode B, usually called a *dynode*, which is more positive with respect to A by 100 volts and more electrons are ejected from it. The electrons ejected from B are then directed towards another dynode C and more electrons are ejected from it. C should be at a positive potential with respect to B. This amplification or multiplication of electrons can be repeated many times within the tube by making the electrons strike on a succession of dynodes which are maintained at increasingly positive potential as shown in the figure. As a result of this multiplication, an avalanche of electrons reaches the collector plate F resulting in a strong current flow in the outer circuit.

A photo-multiplier tube is capable of detecting a light beam of very small intensity and is mostly used in detecting faint light signals of visible, infra-red, ultra-violet and gamma radiation. The main advantage of photo-multiplier tube over the ordinary photo-electric cells lies in the fact that if each electron releases n secondary electrons, then in m stages, n^m electrons reach the collector. For example, if n = 5 and m = 6, then $n^m = 15625$ electrons. Thus there is an enormous gain over the signal obtained from an ordinary photo-electric cell.

Example 12.1. *The work function of potassium is 2.0 eV. When ultraviolet light of wavelength 3500 Å (1 Å = 1 Angstrom unit = 10^{-10} m) falls on a potassium surface, what is the maximum energy in electron volts of the photoelectrons?*

Soln.

$$T_{max}\,(K.E._{max}) = h\nu - h\nu_o$$

Since $h\nu_o$ is already expressed in electron volts, we need to compute the quantum energy $h\nu$ of 3500 A.U. light.

$$h\nu = \frac{hc}{\lambda}$$

$$= \frac{6.63\times10^{-34}\,\text{joules}-\text{sec}\times3\times10^{8}\,\text{m/sec}}{3500\times10^{-10}\,\text{m}}$$

$$= 5.7 \times 10^{-19} \text{ joules}$$

$$= \frac{5.7\times10^{-19}}{1.6\times10^{-19}} \text{ eV} \qquad [1\text{eV} = 1.6 \times 10^{-19} \text{ joules}]$$

$$= 3.6 \text{ eV}.$$

Hence the maximum photo-electron energy

$$T_{max} = h\nu - h\nu_0$$

$$= 3.6 \text{ eV} - 2.0 \text{ eV}$$

$$= 1.6 \text{ eV}.$$

Example 12.2. *Light having a wavelength of 5000 Å falls on a material having a photo-electric work function of 1.90 eV. Find (a) the energy of the photon in eV, (b) the kinetic energy of the most energetic photoelectron in eV and in joules, and (c) the stopping potential.*

Soln.

(a) $h\nu = \frac{hc}{\lambda}$

$$= \frac{6.625\times10^{-34}\times3\times10^{8}}{5000\times10^{-10}} \text{ joules}$$

$$= 3.96 \times 10^{-19} \text{ joules}$$

$$= \frac{3.96\times10^{-19}}{1.6\times10^{-19}} \text{ eV} \qquad [1\text{eV} = 1.6 \times 10^{-19} \text{ joules}]$$

$$= 2.47 \text{ eV}.$$

alternately,

$$h\nu = \frac{12,400}{\lambda} \text{ eV}$$

$$= \frac{12,400}{5,000} \text{ eV}$$

$$= 2.47 \text{ eV}.$$

$$T_{max} = h\nu - w_o$$

$$= 2.47 \text{ eV} - 1.90 \text{ eV}$$

$$= 0.57 \text{ eV}$$

$$= 0.57 \times 1.6 \times 10^{-19} \text{ joules}$$

$$= 0.912 \times 10^{-19} \text{ joules.}$$

$$T_{max} = V_s.e$$

where V_s = stopping potential

e = electronic charge.

$$\therefore \quad V_s = \frac{T_{max}}{e} = \frac{0.57 \text{ eV}}{1 \text{ electronic charge}} = 0.57 \text{ V.}$$

Example 12.3. *What is the threshold wavelength for a tungsten surface whose work function is 4.5 eV.*

Soln.

Here $\quad w_o = 4.5 \text{ eV} = 4.5 \times 1.6 \times 10^{-19}$ joules.

Again $\quad w_o = h\nu_o = h\dfrac{c}{\lambda_o}$

or, $\quad \lambda_o = \dfrac{hc}{w_o}$

$$= \frac{6.625\times10^{-34}\times3\times10^{8}}{4.5\times1.6\times10^{-19}}$$

$$= 2760 \times 10^{-10}\text{m}$$

$$= 2760 \text{ Å}.$$

Alternately $\quad w_o = \dfrac{12{,}400}{\lambda_o}$

$$\text{or,}\quad \lambda_o = \frac{12{,}400}{W_o}$$

$$= \frac{12{,}400}{4.5}$$

$$= 2755\ \mathring{A}.$$

***Example 12.4.** The photo-electric threshold of copper is 3200 Å. If ultra-violet light of wavelength 2500 Å falls on it, find (a) the maximum kinetic energy of the photo-electrons ejected, (b) maximum velocity of the photo-electrons and (c) the value of the work function.*

Soln.

(a) Maximum kinetic energy of photo-electrons is

$$T_{max} = h\nu - h\nu_o$$

$$= hc\left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right)$$

$$= \frac{6.625\times10^{-34}\times3\times10^{8}}{10^{-10}}\left(\frac{1}{2500} - \frac{1}{3200}\right)$$

$$= \frac{6.625\times3\times7}{25\times32}\times10^{-18}\ \text{joules}$$

$$= \frac{6.625\times3\times7\times10^{-18}}{25\times32\times1.6\times10^{-19}}\ \text{eV}$$

$$= 1.087\ \text{eV}.$$

Alternately

$$T_{max} = 12{,}400\left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right)\ \text{eV} \quad \text{where } \lambda \text{ and } \lambda_o \text{ are in A.U.}$$

$$= 12{,}400\left(\frac{1}{2500} - \frac{1}{3200}\right)\text{eV}$$

$= 1.087$ eV.

(b) Maximum velocity of the photo-electrons is given by

$$\frac{1}{2}mv_{max}^2 = T_{max}$$

$$\therefore \quad v_{max} = \sqrt{\frac{2T_{max}}{m}}$$

$$= \sqrt{\frac{2\times6.625\times3\times7\times10^{-18}}{25\times32\times9.1\times10^{-31}}}$$

$$= 6.18 \times 10^5 \text{ m/sec.}$$

(c) Work function $w_o = h\nu_o = \dfrac{hc}{\lambda_o}$

$$= \frac{6.625\times10^{-34}\times3\times10^8}{3200\times10^{-10}} \text{ joules}$$

$$= \frac{6.625\times10^{-34}\times3\times10^8}{3200\times10^{-10}\times1.6\times10^{-19}} \text{ eV}$$

$$= 3.88 \text{ eV.}$$

Alternately, $\quad w_o = \dfrac{12,400}{\lambda_o}$ eV λ_o in A.U.

$$= \frac{12,400}{3,200} \text{ eV}$$

$$= 3.88 \text{ eV.}$$

Example 12.5. *A photo-electric surface has a work function of 4 eV. What is the maximum velocity of the photoelectrons emitted by light of frequency 10^{15} Hertz incident on the surface. $h = 6.6 \times 10^{-34}$ joule-sec.; $e = 1.6 \times 10^{-19}$ coulomb; $m = 9 \times 10^{-31}$ kg.*

Soln.

$$w_o = 4 \text{ eV} = 4 \times 1.6 \times 10^{-19} \text{ joules}$$

$$= 6.4 \times 10^{-19} \text{ joules}$$

$$\frac{1}{2}mv^2 = h\nu - w_o$$

$$= 6.6 \times 10^{-34} \times 10^{15} - 6.4 \times 10^{-19}$$

$$= 0.2 \times 10^{-19} \text{ joules}$$

$$\therefore \quad v = \sqrt{\frac{2\times0.2\times10^{-19}}{m}}$$

$$= \sqrt{\frac{0.4\times10^{-19}}{9\times10^{-31}}}$$

$$= 2.107 \times 10^5 \text{ m/sec.}$$

Example 12.6. *Calculate the threshold frequency and the corresponding wavelength of radiation incident on a certain metal whose work function is* 3.31×10^{-19} *J. Given, Planck's constant =* 6.62×10^{-34} *J-s.*

Soln.

Work function, $w_o = h\nu_o$

$$\text{or,} \quad \nu_o = \frac{w_0}{h}$$

Here, $w_o = 3.31 \times 10^{-19}$ J.

$h = 6.62 \times 10^{-34}$ J-s

$$\therefore \quad \nu_o = \frac{3.31\times10^{-19}\text{ J}}{6.62\times10^{-34}\text{ J-s}}$$

$$= 5 \times 10^{14} \text{ Hertz}$$

$$\therefore \quad \lambda_o = \frac{c}{\nu_0} = \frac{3\times10^8 \text{ m/sec}}{5\times10^{14}}$$

$$= 6 \times 10^{-7} \text{ m} = 6000 \overset{o}{\text{A}}.$$

Example 12.7. *Photo-electrons are emitted with a maximum speed of 7×10^5 m/sec from a metal surface when light of frequency 8×10^{11} kHz falls on it. What is the threshold frequency of the metal.*

Soln.

$$T_{max} = h\nu - h\nu_o$$

$$\frac{1}{2}mv^2 = h(\nu - \nu_o)$$

or, $$\frac{1}{2} \times 9.1 \times 10^{-31} \times (7 \times 10^5)^2 = 6.6 \times 10^{-34}(8 \times 10^{14} - \nu_o)$$

or, $$6.6 \times 10^{-34}\nu_o = (52.8 - 22.3) \times 10^{-20}$$

or, $$\nu_o = \frac{30.5 \times 10^{-20}}{6.6 \times 10^{-34}} = 4.62 \times 10^{14} \text{ Hz.}$$

Example 12.8. *A tungsten cathode whose threshold wavelength is 2300 Å is irradiated by ultraviolet light of wavelength 1800 Å. Calculate (i) the work function for tungsten and (ii) the maximum energy of the photoelectrons emitted, both in electron-volts.*

Soln.

(i) the work function, $w_o = \dfrac{12{,}400}{\lambda_o}$ eV.

where λ_o is expressed in Angstrom unit.

$$\therefore \quad w_o = \frac{12{,}400}{3{,}200} = 5.4 \text{ eV.}$$

(ii) Similarly

$$T_{max} = 12400\left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right) \text{eV}$$

where λ and λ_o are in Angstrom units, .

$$\therefore \quad T_{max} = 12400\left(\frac{1}{1800} - \frac{1}{2300}\right)$$

$$= 1.5 \text{ eV.}$$

Example 12.9. *Light of wavelength 4300 Å is incident on (a) nickel surface of work function 5 electron volts and (b) potassium surface of work function 2.3 electron volts. Find out, if electrons will be emitted, and if so, the maximum velocity of the emitted electrons in each case.*

Soln.

(i) For the nickel surface

$w_o = h\nu_o = 5$ eV

$= 5 \times 1.6 \times 10^{-19}$ J

$$\therefore \quad \nu_o = \frac{w_0}{h} = \frac{5\times1.6\times10^{-16}}{6.625\times10^{-34}} \text{ Hz}$$

$$\therefore \quad \lambda_o = \frac{c}{\nu_0} = \frac{3\times10^8\times6.624\times10^{-34}}{5\times1.6\times10^{-19}} \text{ m}$$

$= 2484 \times 10^{-10}$ m

$= 2484$ Å

2484 Å is the longest wavelength at which photo-electric emission will occur at nickel surface. Since λ_o is less than the wavelength of the incident radiation ($\lambda = 4300$ Å), electrons will not be emitted from nickel surface.

(ii) For the potassium surface

$w_o = 2.3$ electron volts

$= 2.3 \times 1.6 \times 10^{-19}$ J.

$$\therefore \quad \lambda_o = \frac{c}{\nu_0} = \frac{ch}{w_0} = \frac{3\times10^8\times6.624\times10^{-34}}{2.3\times1.6\times10^{-19}} \text{ m.}$$

$= 4389 \times 10^{-10}$ m.

$= 4389$ Å.

As λ_o is greater than λ (= 4300 $\overset{o}{A}$), electrons will be emitted.

The maximum energy of the emitted electrons

$$\frac{1}{2}mv^2_{max} = h\nu - h\nu_o$$

The energy $h\nu$ corresponding to $\lambda = 4300\ \overset{o}{A}$

$$= \frac{12400}{\lambda}\ eV = \frac{12400}{4300}\ eV = 2.9\ eV.$$

$$\therefore \quad \frac{1}{2}mv^2_{max} = (2.9 - 2.3)\ eV$$

$$= 0.6\ eV = 0.6 \times 1.6 \times 10^{-19}\ J.$$

$$\text{or,} \quad v_{max} = \sqrt{\frac{2\times0.6\times1.6\times10^{-19}}{m}}$$

$$= \sqrt{\frac{2\times0.6\times1.6\times10^{-19}}{9.1\times10^{-31}}}$$

$$= \sqrt{0.21\times10^{12}} = 4.6 \times 10^5\ m/s.$$

Example 12.10. *Radiant energy of wavelength 6600 $\overset{o}{A}$ is falling on a layer of barium oxide at a rate of 3 x 10^{-8} watt. Calculate the number of photo-electrons yielded, assuming that every photon incident on the barium oxide layer produces photo-electric effect. Calculate also the maximum velocity of photo-electrons released by the radiation (work function of barium = 1.4 eV).*

Soln.

Rate at which the energy is falling on the barium layer

$= 3 \times 10^{-8}$ watt

$= 0.3 \times 10^{-7}$ joules/sec.

Suppose the number of electrons emitted per second is n when an energy at the rate of 0.3×10^{-7} joules/sec falls on the barium layer.

$\therefore$ energy necessary for ejection of one electron = $\frac{0.3\times10^{-7}}{n}$ joule

Now $E = h\nu = hc/\lambda$

$$\therefore \quad \frac{0.3}{n}\times10^{-7} = \frac{6.625\times10^{-34}\times3\times10^{8}}{6600\times10^{-10}}$$

$$\text{or,} \quad n = \frac{0.3\times66}{6.625\times3}\times10^{11}$$

$$= 10^{11} \text{ electrons.}$$

Number of electrons yielded per sec = 10^{11}

The maximum velocity of the photo-electrons is given by

$$\frac{1}{2}mv^2_{max} = h\nu - w_0$$

$$= \frac{hc}{\lambda} - w_0 \quad \text{where } w_0 = 1.4 \text{ eV}$$

$$= 1.4\times1.6\times10^{-19}\text{ J}$$

$$= \frac{6.625\times10^{-34}\times3\times10^{8}}{6600\times10^{-10}} - 1.4\times1.6\times10^{-19}$$

$$= \frac{6.625\times3}{66}\times10^{-18} - 1.4\times0.16\times10^{-18}$$

$$= 0.077\times10^{-18}\text{ Joules.}$$

$$\therefore \quad v_{max} = \sqrt{\frac{2\times0.077\times10^{-18}}{m}}$$

$$= \sqrt{\frac{2\times0.077\times10^{-18}}{9.1\times10^{-31}}}$$

$$= \left[\frac{2\times0.77}{9.1}\right]^{1/2}\times10^{6}$$

$$= 4.1\times10^{5}\text{ m/sec.}$$

Example 12.11. *A surface having work function 1.51 eV is illuminated by light of wavelength 4000 Å. Calculate (i) the maximum kinetic energy of the ejected electrons and (ii) the stopping potential.*

Soln.

The energy content of a photon having a wavelength of 4000 Å

$$= \frac{12{,}400}{4{,}000} \text{ eV}$$

$$= 3.1 \text{ eV.}$$

Maximum kinetic energy of electrons ejected by photons of wavelength 4000 Å

$$= h\nu - w_o$$

$$= 3.1 - 1.51 \qquad \text{(here } w_o = 1.51 \text{ eV)}$$

$$= 1.59 \text{ eV}$$

$$= 1.59 \times 1.6 \times 10^{-19} \text{ joules}$$

$$= 2.544 \times 10^{-19} \text{ joules.}$$

Also,

$$\text{K. E.}_{max} = h\nu - w_o = e.V_s$$

where V_s is the stopping potential in volts for these electrons

$$\therefore \quad e.V_s = 1.59 \text{ eV}$$

$$\text{or,} \quad V_s = \frac{1.59 \times 1.6 \times 10^{-19} \text{ J}}{e}$$

$$= \frac{1.59 \times 1.6 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ V}$$

$$= 1.59 \text{ V.}$$

$$\left[\text{or, } V_s = \frac{1.59 \text{ eV}}{e} = 1.59 \text{ V}\right]$$

Example 12.12. *The stopping potential for electrons emitted from a metal due to photo-electric effect is found to be 1 V for light of 2500 Å. Calculate the work function of the metal in eV.*

Soln.

Maximum energy of the ejected photo-electron

$$E_{max} = e.V_s \text{ joules}$$

$$= V_s \text{ electron-volt}$$

where V_s is the stopping potential in volt and e is the electronic charge (= 1.6×10^{-19} C)

Hence, E_{max} in this case

$= V_s$

$= 1$ eV.

Energy content of photon of wavelength 2500 Å.

$$= \frac{12{,}400}{2{,}500} \text{ eV}$$

$= 4.96$ eV.

Hence, from $E_{max} = h\nu - w_o$, we have

$$w_o = h\nu - E_{max}$$

where w_o is the required work function.

$\therefore \quad w_o = 4.96 - 1 = 3.96$ eV.

Example 12.13. *A photon of wavelength 3310 Å falling on a photo-cathode ejects an electron of energy 3×10^{-19} J and one of wavelength 5000 Å ejects an electron of energy 0.972×10^{-19} J. Calculate the value of Planck's constant and the threshold wavelength for the photo cathode.*

Soln.

$$h\nu = w_o + T$$

for both cases, threashold wavelength/is same.
work function

or, $\frac{hc}{\lambda} = w_o + T$

In the first case,

$$\frac{h\times3\times10^{8}}{3310\times10^{-10}} = w_o + 3 \times 10^{-19} \quad \text{(i)}$$

In the second case,

$$\frac{h\times3\times10^{8}}{5000\times10^{-10}} = w_o + 0.972 \times 10^{-19} \quad \text{(ii)}$$

Subtracting (ii) from (i),

$$h \times 3 \times 10^{18}\left[\frac{1}{3310}-\frac{1}{5000}\right] = [3.0 - 0.972] \times 10^{-19}$$

or, $\frac{h\times3\times10^{18}\times1690}{1655\times10^{4}} = 2.028 \times 10^{-19}$

or, $h \times 0.306 \times 10^{15} = 2.028 \times 10^{-19}$

or, $h = \frac{2.028\times10^{-19}}{0.306\times10^{15}} = 6.627 \times 10^{-34}$ J-sec.

Substituting this value of h in (ii), we get

$$\frac{6.627\times10^{-34}\times3\times10^{8}}{5000\times10^{-10}} = w_o + 0.972 \times 10^{-19}$$

or, $3.976 \times 10^{-19} = w_o + 0.972 \times 10^{-19}$

or, $w_o = [3.976 - 0.972] \times 10^{-19}$

$= 3.002 \times 10^{-19}$

From $w_o = h\nu_o = \frac{ch}{\lambda_o}$, we have

$$\lambda_o = \frac{ch}{w_o} = \frac{3\times10^{8}\times6.627\times10^{-34}}{3.002\times10^{-19}}$$

$= 6.620 \times 10^{-7}$ m

$= 6620 \times 10^{-10}$ m

$= 6620$ Å.

Example 12.14. *The stopping potential is 4.6 V for light of frequency 2 x 10^15 Hz. When light of frequency 4 x 10^15 Hz is used, the stopping potential is 12.9 V. Calculate the value of Planck's constant.*

Soln.

$$h\nu = w_o + T_{max}$$

$$= h\nu_o + e.V_s$$

where ν_o is the threshold frequency and V_s is the stopping potential.

or, $e.V_s = h(\nu - \nu_o)$

In the first case,

$$4.6\ e = h\ (2 \times 10^{15} - \nu_o) \quad \text{(i)}$$

In the second case,

$$12.9\ e = h\ (4 \times 10^{15} - \nu_o) \quad \text{(ii)}$$

Subtracting (i) from (ii)

$$8.3\ e = h \cdot 2 \times 10^{15}$$

Substituting $e = 1.6 \times 10^{-19}$ C, we get

$$h \cdot 2 \times 10^{15} = 8.3 \times 1.6 \times 10^{-19}$$

or, $$h = \frac{8.3 \times 1.6 \times 10^{-19}}{2 \times 10^{15}}$$

$$= 6.64 \times 10^{-34} \text{ J-s.}$$

Example 12.15. *A certain metallic surface is illuminated by monochromatic light of variable wavelength. No photoelectrons are emitted above a wavelength of 5000 Å. With an unknown wavelength, a stopping potential of 3.1 V is necessary to stop photo-electric current. Find the unknown wavelength.*

Soln.

Kinetic energy of the emitted electrons,

$$T = 12{,}400\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)\text{eV where } \lambda \text{ and } \lambda_o \text{ are in } \text{Å}.$$

If V_s is the stopping potential in volts, then the kinetic energy of the emitted electrons is V_s electron volt.

$$3.1 = 12{,}400\left(\frac{1}{\lambda}-\frac{1}{5000}\right)$$

$$= \frac{12400}{\lambda} - 2.48$$

or, $$\frac{12400}{\lambda} = 3.1 + 2.48$$

or, $$5.58\,\lambda = 12400$$

or, $$\lambda = \frac{12400}{5.58} = 2222\ \text{Å}.$$

Example 12.16. *If the photo-electric threshold of metallic silver is 3800 Å and ultraviolet light of λ = 2600 Å falls on it, find (a) maximum kinetic energy of photo-electron ejected (b) maximum velocity of photo-electrons and (c) value of work function in joules.*

(c) $$w_o = \frac{12400}{3800}$$

$$= 3.26\ \text{eV}.$$

(a) Energy of photon of wavelength 2600 Å

$$= \frac{12400}{2600} = 4.77 \text{ eV}$$

$$\therefore \quad T_{max} = h\nu - w_o$$

$$= 4.77 - 3.26$$

$$= 1.51 \text{ eV}$$

$$= 1.51 \times 1.6 \times 10^{-19} \text{ J.}$$

(b) $\frac{1}{2}mv_{max}^2 = T_{max}$

$$\frac{1}{2} \times 9.1 \times 10^{-31} \times (v_{max})^2 = 1.51 \times 1.6 \times 10^{-19}$$

or, $v_{max} = \sqrt{\frac{2\times1.51\times1.6\times10^{-19}}{9.1\times10^{-31}}}$

$$= 7.28 \times 10^5 \text{ m/sec.}$$

Example 12.17. *A certain metal has a threshold wavelength of 6000 Å. Find the stopping potential when the metal is irradiated with,*

(i) monochromatic light of wavelength 4000 Å

(ii) light having twice the intensity and frequency and

(iii) if a material having double the work function were used, what would be the answers to (i) and (ii) above?

Soln.

(i) The maximum kinetic energy of photo-electrons is given by

$$T_{max} = e.V_s = h\nu - h\nu_o$$

where V_s is the stopping potential

$$eV_s = h\nu\left(\frac{1}{\lambda}-\frac{1}{\lambda_o}\right)$$

$$= 6.625 \times 10^{-34} \times 3 \times 10^{8}\left(\frac{1}{4000}-\frac{1}{6000}\right) \times 10^{10}$$

$$= \frac{6.625 \times 3 \times 2000 \times 10^{-16}}{4000 \times 6000} \text{ joule}$$

$$= \frac{6.625 \times 3 \times 2 \times 10^{-19}}{24 \times 1.6 \times 10^{-19}} \text{ eV}$$

$$= 1.03 \text{ eV.}$$

∴ stopping potential, $V_s = 1.03$ volts.

(ii) Increase of intensity will have no effect on the stopping potential.

When the frequency of light is twice, the wavelength of light becomes half of that in (i) *i.e.*, $\lambda = 4000/2 = 2000\,\mathring{A}$. So the stopping potential is given by

$$eV_s = 6.625 \times 10^{-34} \times 3 \times 10^{8}\left(\frac{1}{2000}-\frac{1}{6000}\right) \times 10^{10}$$

$$= \frac{6.625 \times 3 \times 10^{-16} \times 4000}{2000 \times 6000} \text{ joules}$$

$$= \frac{6.625 \times 12 \times 10^{-16}}{12 \times 1.69 \times 10^{-19}} \text{ eV}$$

$$= 4.14 \text{ eV}$$

∴ $V_s = 4.14$ volts.

(iii) If the work function is doubled, then λ_o is reduced to half *i.e.*, $\lambda_o = 6000/2 = 3000\,\mathring{A}$. Since the wavelength of the incident light is $4000\,\mathring{A}$, it would not be able to produce photo-emission.

In the second case, $\lambda = 2000\ \mathring{A}$.

Hence

$$e.V_s = 12400\left(\frac{1}{\lambda} - \frac{1}{\lambda_o}\right) eV$$

$$= 12400\left(\frac{1}{2000} - \frac{1}{3000}\right)$$

$$= \frac{12400 \times 1000}{2000 \times 3000}\ eV$$

$$= 2.06\ eV$$

Hence the stopping potential,

$$V_s = 2.06 \text{ volts.}$$

Example 12.18. *It takes 4.2 eV to remove one of the least tightly bound electrons from a metal surface. When ultra-violet photons of a single frequency strike a metal, electrons with kinetic energies from zero to 2.6 eV are ejected. What are the energy and wavelength of the incident photons?*

Soln.

$w_o = 4.2$ eV. $\qquad T_{max} = 2.6$ eV.

We have

$$E_{photon} = h\nu = T_{max} + w_o$$

$$= 2.6 + 4.2 = 6.8 \text{ eV.}$$

Also

$$E_{photon} = \frac{12400}{\lambda}\ eV$$

or, $$\lambda = \frac{12400}{6.8}\ \mathring{A}$$

$$= 1823\ \mathring{A}.$$

Example 12.19. *When violet light of λ = 4000A.U. strikes the cathode of a photo-cell, a retarding potential of 0.4 eV is required to stop the emission of electrons. Find (a) the frequency of the light, (b) the energy of the wavelength, (c) work function of the surface, threshold frequency and its wavelength, the net energy and velocity with which the electron leaves the surface.*

Soln.

(a) $\nu = \frac{c}{\lambda} = \frac{3\times10^{8}}{4000\times10^{-10}} = 7.5 \times 10^{14}$ Hz.

(b) $E = h\nu = 6.6925 \times 10^{-34} \times 7.5 \times 10^{14}$ joules.

$$= \frac{6.625\times10^{-34}\times7.5\times10^{14}}{1.6\times10^{-19}} \text{ eV}$$

$= 3.1$ eV.

Work function, $w_o = h\nu_o = h\nu - \frac{1}{2}mv^2$

$$= h\nu - eV_s \qquad (\frac{1}{2}mv^2 = e.V_s)$$

$$= 3.1 - 0.4 = 2.7 \text{ eV}.$$

Threshold frequency, $\nu_o = \frac{w_o}{h} = \frac{2.7\times1.6\times10^{-19}}{6.624\times10^{-34}}$

$$= 6.53 \times 10^{14} \text{ c/s (Hz)}$$

Corresponding wavelength, $\lambda = \frac{c}{\nu_o} = \frac{3\times10^{8}}{6.53\times10^{14}}$

$$= 4590 \times 10^{-10} \text{ m}$$

$$= 4590 \text{ A.U.}$$

Energy of the emitted electron,

$$E_{max} = h\nu - h\nu_o$$

$$= 3.1 - 2.7 = 0.4 \text{ eV.}$$

Velocity of the emitted electron,

$$v = \sqrt{\frac{2E}{m}} = 1.2 \times 10^8 \text{ m/sec.}$$

Example 12.20. *Compute the number of photons of yellow light of wavelength 6000 Å required to make an erg of energy.*

Soln.

Energy content of a photon,

$$E = h\nu = h.\frac{c}{\lambda} = \frac{6.625\times10^{-34}\times3\times10^{8}}{6000\times10^{-10}} \text{ J}$$

$$= \frac{6.625\times10^{-34}\times3\times10^{8}\times10^{7}}{6000\times10^{-10}} \text{ ergs}$$

$$= \frac{6.625\times3\times10^{-19}}{6\times10^{-7}}$$

$$\text{No. of photons} = \frac{1\times6\times10^{-7}}{6.625\times3\times10^{-19}}$$

$$= \frac{6\times10^{-7}}{20\times10^{-19}}$$

$$\simeq 3 \times 10^{-7} \times 10^{18}$$

$$\simeq 3 \times 10^{11}.$$

EXERCISE

1. What is photo-electric effect? Describe an experiment for studying the phenomenon of photo-electric emission and discuss the results obtained therefrom. How has this phenomenon been explained by Einstein?

2. What is photo-electric effect? Define photo-electric work function and threshold frequency.

3. Give an account of the observed facts about photo-electric emission and indicate in a general way the difficulties encountered by the classical wave theory of light in explaining these facts.

4. What was the hypothesis on which Einstein based his photo-electric equation? Discuss how this equation provides a satisfactory explanation of the observed facts regarding photo-electric emission.

5. State and explain the laws governing photo-electric emission. Establish Einstein's photo-electric equation and show how it explains these laws.

6. Establish Einstein's photo-electric equation and show that the maximum velocity of the emitted electrons depends upon the frequency of the incident radiation and not on its intensity.

7. What is meant by stopping potential in connection with photo-electric effect? Show that the stopping potential varies linearly with the frequency of the incident radiation but is independent of its intensity.

8. Establish Einstein's photo-electric equation. Describe Millikan's method of verifying this equation.

9. Describe Millikan's experiment and show how Einstein's photo-electric equation can be verified. How would you determine the value of (i) Planck's constant, (ii) the threshold frequency and (iii) the work function by Millikan's experiment?

10. Establish Einstein's photo-electric equation. Describe an experiment to verify this equation.

11. Describe the construction, working and applications of various types of photo-cells.

12. Describe the construction and working of a photo-voltaic cell.

13. Discuss the characteristics of photo-electric emission. Establish Einstein's photo-electric equation and show how it can explain these characteristics.

14. What is photo-electric effect? State and explain the laws of photo-electricity.

15. State the laws of photo-electricity and explain these laws with the help of Einstein's theory of photo-electricity.

16. Give quantum interpretation of photo-electric emission indicating briefly the failure of classical theory and the success of the quantum theory to explain the results of photo-electric emission.

17. Find the energy content of ultra-violet light photons of wavelength 3000 A.U. in (i) joules and (ii) electron-volt. **[(i) 6.62×10^{-19} J, (ii) 4.13 eV]**

18. Calculate the frequency and amount of energy (in eV) associated with a quantum of electro – magnetic radiation at each of the following wavelengths (i) 1 m (ii) 1 mm (iii) 5000 Å and (iv) 1 Å. **[(i) 3×10^8 Hz; 1.25×10^{-16} eV, (ii) 3×10^{11} Hz; 1.25×10^{-3} eV, (iii) 6×10^{14} Hz; 1.25 eV, (iv) 3×10^{18} Hz; 1.25×10^4 eV]**

19. Evaluate the threshold wavelength of photo-electric material whose work function is 2 eV. **[6.21×10^{-7} m]**

20. The photo-electric threshold of copper is 3200 Å. If the ultra-violet light of wavelength 2500 Å falls on it, find (a) value of work function of copper (b) maximum kinetic energy of the ejected photo-electrons. **[(a) 6.21×10^{-19} J, (b) 1.739×10^{-19} J]**

21. Silver has a photo-electric function of 4.5 eV. Find (i) its threshold wavelength, (ii) maximum velocity of ejected photo-electrons with ultra-violet light of wavelength 2000 Å, (iii) stopping potential for the most energetic photo-electrons. **[(i) 2760 Å, (ii) 7.71×10^5 m/sec, (iii) 1.7 V]**

22. The wavelength of the photo-electric threshold for silver is 3250 Å. Determine the maximum energy of the electrons ejected by light of wavelength 2537 Å. **[1.625×10^{-19} J]**

23. The long wavelength limit for photo-electric effect for a certain metal is 2750 Å. Find (i) work function w_o (ii) the maximum velocity of electrons ejected from the metal by radiation of $\lambda = 1800$ Å (iii) maximum kinetic energy of these electrons **[(i) 4.5 eV (ii) 9.1×10^5 m/s (iii) 3.8×10^{-19} J]**

24. Calculate the frequency of the incident radiation which eject from a metal surface electrons fully retarded by a reverse potential of 3 V. The threshold frequency for this metal is 6×10^{14} Hz. Also find the work function of this metal. **[13.2×10^{14} Hz; 2.48 eV]**

25. With what velocity must an electron travel so that its kinetic energy is equal to the energy of a photon of wavelength $\lambda = 5200 \text{ Å}$? **[9.2 x 10^5 m/s]**

26. Light of $\lambda = 5000 \text{ Å}$ falls on a photo-metal having work function of 1.9 eV. Find (a) energy of the photon in eV (b) K.E. of the most energetic photo-electron in joule and eV (c) the stopping potential. **[(a) 2.47 eV (b) 0.57 eV (c) 0.57 V]**

27. If the photo-electric threshold of a given metal lies in the green of the visible spectrum $\lambda = 5000 \text{ Å}$, and ultra-violet light of wavelength $\lambda = 4000 \text{ Å}$ falls on it, find (a) K.E. of ejected photo-electrons (b) velocity of the photo-electrons and (c) the value of the work function. **[(a) 9.8 x 10^{-20} joules (b) 4.69 x 10^5 m/sec (c) 0.2 x 10^{-19} joule]**

28. When light of frequency 1.3 x 10^{15} c/s falls on a metal, photo-electrons are emitted with a maximum energy of 1.8 eV. Find the work function of the metal in ergs and eV. Also find the threshold frequency for the metal. **[w_o = 3.58 eV, 5.728 x 10^{-12} ergs, ν_o = 8.05 x 10^{14} cycles]**

29. A metal surface when illuminated by light of frequency 0.90 x 10^{15} cycles/sec emits electrons which can be stopped by a retarding potential of 0.60 volt. When the same surface is illuminated by light of frequency 1.26 x 10^{15} cycles/sec, the required retarding potential is 2.1 volts. Deduce the value of Planck's constant and work function of the metal. **[6.67 x 10^{-34} Joule-sec, 3.31 x 10^{-19} J]**

30. A certain metal has a threshold wavelength of 6526 Å. Find the stopping potential when the metal is irradiated with

    (a) monochromatic light of wavelength 3263 Å

    (b) monochromatic light having twice the frequency and three times the intensity of that in (a).

    [(a) 1.9 volt (b) 5.7 volt; no change with intensity]

31. A radiation of frequency 10^{15} Hz falls on a photo-cathode and eject electrons with maximum energy of 4.2 x 10^{-19} joule. If the frequency of radiation is changed to 5 x 10^{14} Hz, the maximum energy of ejected electrons becomes 0.9 x 10^{-19} Joule. Calculate the value of Planck's constant threshold frequency and work function of photo-cathode material. **[6.6 x 10^{-34} J-sec; 3.64 x 10^{14} Hz; 2.4 x 10^{-19} J]**

32. The photo-electric threshold for tungsten is 2300 A.U. Find the energy of the emitted electrons from the surface by ultra-violet light of λ = 180 A.U. **[63.58 eV]**

33. When a tungsten surface is illuminated by light of wavelength 1800 A.U. the maximum energy of the electrons liberated is 1.50 eV. Photo-electric emission for tungsten ceases when the wavelength of the incident light exceed 2300 $\mathring{A}$. What is the magnitude of Planck's constant? **[6.624 x 10^{-34} J-sec]**

34. Ultra-violet light of wavelength 3000 x 10^{-8} cm falls on the surface of a metal of work function 2.28 eV and ejects an electron. What will be the velocity of the ejected electron? (Mass of electron = 9.1 x 10^{-28} gm, 1 eV = 1.6 x 10^{-12} erg and h = 6.6 x 10^{-27} erg-sec). **[2.55 x 10^8 cm/sec]**

35. Find the maximum velocity of the photo-electrons when light of λ = 2.537 x 10^{-7} m falls on a metal whose work function is 4.51 eV. **[3.66 x 10^5 m/sec]**

36. When light of frequency 5 x 10^{14} Hz falls on a metal surface, electrons of energy 2.31 x 10^{-19} J are emitted. Find the wavelength of ultra-violet light which falling on the same surface would liberate electrons of energy 8.93 x 10^{-19} J. (c = 3 x 10^8 m/sec, h = 6.62 x 10^{-34} J-s) **[2 x 10^{-7} m]**

37. Radiation of wavelength 5000 $\mathring{A}$ and intensity 2 x 10^{-2} watt/sq.cm. falls on a photo-sensitive surface. Assuming that every absorbed photon results in the ejection of a photo-electron, find how many photo-electrons are produced per square centimeter per second. **[5 x 10^{16} per sq.cm]**

38. A retarding potential of 0.4 eV is required to block the movement of electrons from the cathode of a photo-cell when violet light of λ = 4000 $\mathring{A}$ strikes the surface. Find (a) the frequency of this light, (b) the energy of this λ, (c) the work function of this surface, (d) the threshold frequency, (e) λ for this frequency, (f) energy for λ = 3000 $\mathring{A}$, (g) net energy after the electron leaves the metal surface, (h) energy in joule, (i) the velocity of this electron after it leaves the surface. (h = 6.63 x 10^{-34} J-sec = 4.13 x 10^{-15} eV)

[(a) 7.5 x 10^{14} Hz (b) 3.09 eV (c) 2.7 eV (d) 6.53 x 10^{14} Hz (e) 4593 $\mathring{A}$ (f) 4.13 eV (g) 1.43 eV (h) 2.29 x 10^{-12} J (i) 7.1 x 10^5 m/sec]

39. Find (a) the longest λ that can emit photo-electrons from a caesium surface for which w_0 = 1.35 eV, (b) the maximum velocity of the emitted photo-electron when illuminated with light of $\lambda = 4 \times 10^{-7}$ m, and (c) the potential difference just sufficient to prevent a current from passing through a caesium photo-cell when illuminated with λ = 4×10^{-7} m. **[(a) 9×10^{-7} m, (b) 7.9×10^{5} m/sec (c) 1.7 volts]**

40. Calculate the longest wavelength of the incident radiation which will eject electrons from a metal whose work function as 6 eV. Planck's constant h = 6.62×10^{-34} J-s. **[2070 Å]**

41. Calculate in electron-volt, the energy of a quantum of light of wavelength = 5.3×10^{-7}m. **[2.34 eV]**

42. Photo-electrons with a velocity of 4×10^{5} m/sec are emitted from a metal surface when light of wavelength 6000 Å falls on it. What is the photo-electric threshold wavelength for the metal? **[7631 Å]**

43. A photon of wavelength 3210 Å falls on a photo-cathode and an electron of energy 3×10^{-19} joule is ejected; if the wavelength of the incident photon is changed to 5000 Å, the energy of the ejected electron is 0.972×10^{-19} joule. Calculate the value of the Planck's constant and the threshold wavelength of the photo-cathode (c = 3×10^{8} m/s) **[6.62×10^{-34} J.s; 6620 Å]**

44. When light of wavelength l = 5500 Å falls on a surface, photoelectrons are ejected for which the stopping potential is V_S = 0.19 volt. If the wavelength of the incident photon is changed to 1900 Å, calculate (a) the stopping potential (b) the work junction of the surface. **[4.47 v; 2.07 eV; 5×10^{14} Hz]**

45. The work function of molybdenum is 4.15 volts. If ultra-violet light of wavelength 1000 Å is incident upon molybdenum, find maximum velocity of the ejected photo-electrons. **[1.72×10^{6} m/sec.]**

CHAPTER XIII

X - RAYS

The photo-electric effect provides convincing evidence that light photons can transfer energy to electrons. Is the inverse process also possible? That is can the kinetic energy of a moving electron be partly or wholly converted into a photon? As it happens, the inverse photo-electric effect not only does occur, but had already been discovered – though not at all understood, before Planck put forward his quantum theory.

The history of science has many instances of *accidental* discovery and of these the discovery of X-rays is a prime example. In 1895 Roentgen was studying the phenomenon of gaseous discharge. When a high voltage of several tens of kilovolts was passed though a cathode-ray tube, exhausted to 10^{-3} mm of Hg, Roentgen noticed that a screen coated with barium platinocyanide, present in the laboratory at a distance from the cathode-ray tube fluoresced brilliantly. This effect persisted even when the tube was wrapped with a layer of black card-board. Roentgen soon established that the agency responsible for the fluorescence originated at the point at which the stream of energetic electrons struck the glass wall of the tube. Because of its unknown nature, he gave this agency the name X-rays. Roentgen found that X-rays could manifest themselves by darkening wrapped photographic plates, discharging charged electroscopes, as well as by causing fluorescence in a number of different substances. He further observed that X-rays can penetrate considerable thicknesses of materials of low atomic number, whereas substances of high atomic number are relatively opaque.

There is probably no man made discovery that attracted public attention more quickly than the discovery of X-rays. The fact that the rays permitted one to *see* through opaque objects was sensational material for *press* and there was great consternation lest, by their use, fully dressed people might be made to appear unclothed. However, when such speculation died down, the value of X-rays in setting broken bones was widely appreciated, and the rays were quickly put to this use by surgeons only a few month's after Roentgen's discovery.

13.1 Nature of X-rays

The discovery of X-rays aroused the interest of physicists and many joined in the investigation of their properties. Not long after their discovery it began to be suspected that X-rays are electromagnetic waves. After all, according to electron-magnetic theory, an accelerated electric charge will radiate electromagnetic waves. A rapidly moving electron suddenly brought to rest is certainly accelerated and the radiation produced under these circumstances is given the German name *bremsstrahlung* (braking radiation). Haga and Wind, in 1899, showed that X-rays are a wave motion phenomenon and estimated their wavelength to be $\lambda \sim 10^{-8}$ cm. The wave nature of X-rays was established in 1906 by Barkla who showed that they can be polarized and that the waves are transverse in nature. Max von Laue, in 1912, devised an experiment for actual measurement of the wavelength of X-rays, by suggesting that a crystal can act as a diffraction grating. This suggestion was first put into practice by Friedrick and Knipping and the technique was subsequently perfected by the Braggs. The wavelengths measured in these experiments ranged from 1.3×10^{-11} to 4.8×10^{-11}m (0.13 Å to 0.48 Å). The wavelengths are thus 10^{-4} of those in visible light, and hence having quanta 10^{4} times as energetic. For purposes of classification, electromagnetic radiations with wavelengths in the approximate interval from 10^{-11} to 10^{-8}m (0.1 Å to 100 Å) are today considered as X-rays.

Properties of X-rays

1. X-rays are electromagnetic radiations of the same nature as light, of shorter wavelength of the order of 10^{-10}m and are photons of high energy.

2. They are not deflected by a magnetic or electric field; hence they are not charged particles and differ from cathode rays.

3. They cannot be reflected or refracted by ordinary methods.

4. They affect photographic plates

5. They travel with the same velocity as that of light *i.e.*, 3 x 10^8 m/sec.

6. They produce photoelectric effect; for they liberate electrons when they fall on certain substances.

7. They exhibit the phenomena of interference, diffraction and polarization. In the respects mentioned above, X-rays more or less resemble light waves.

8. They are not refracted on passing from one medium to the other, thus differing from ordinary light waves. The absence of any perceptible X-ray refraction could be attributed to very short wavelengths of X-rays, below those in the ultra-violet range. The refractive index of a substance decreases to unity (corresponding to straight-line propagation) with decreasing wavelength.

9. They ionize gas through which they pass.

10. They excite fluorescence on barium-platino cyanide, zinc sulphide and cadmium tungstate.

11. They are highly penetrating radiation. They can pass through many solids. Their penetrating power depends on their frequency; the higher the frequency the greater the penetrating power.

12. When very high velocity electrons fall on a target, two kinds of X-rays are produced. (a) continuous radiation and (b) characteristic radiations. Characteristic radiation depends on the nature of the target.

13.2 Production of X-rays

(i) Gas-filled X-rays tube : Fig. 13.1 represents an early form of an X-ray tube similar to that used by Roentgen. It consists of a spherical glass tube G exhausted to a pressure of 10^{-3} mm of Hg. C is a concave cathode made of aluminium. The cathode rays from the concave cathode are focussed on a metal target T. Tungsten is used as

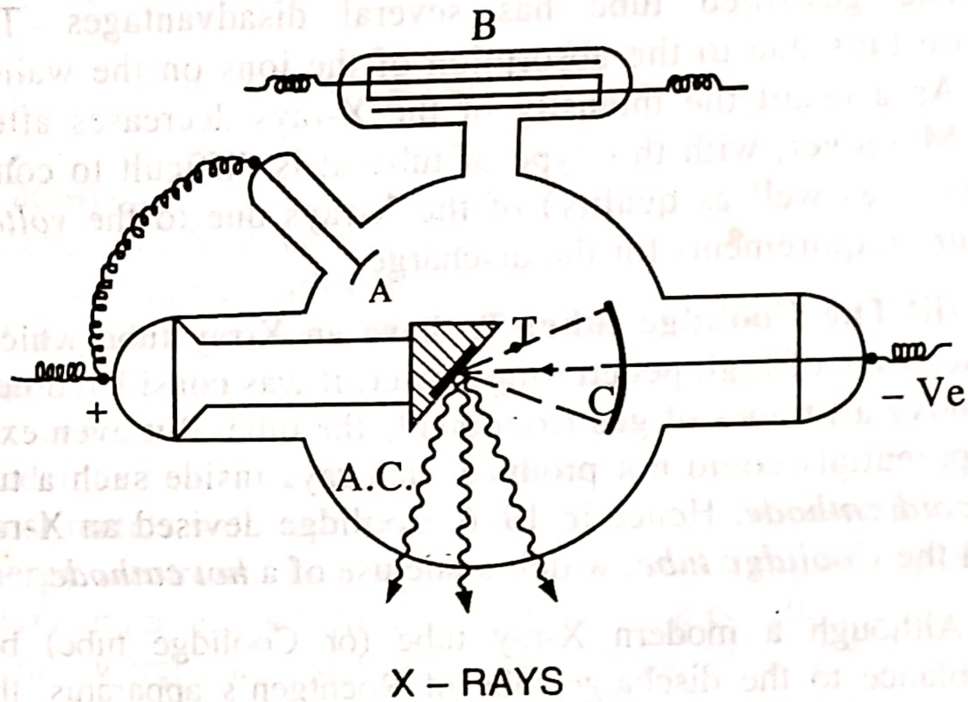


Fig. 13.1

the target because of its high melting point. The striking face of the target is kept at 45° with the incident beam of the cathode rays. The target is surrounded by a massive block of copper. The target and the copper block together is called the anticathode (A.C.). A separate anode A is kept in a side tube and is connected to the anti-cathode externally. The presence of a separate anode makes the discharge steady. B is occlusion tube containing platinized asbestos in which gases have been occluded. When the X-ray tube stops working due to very low pressure in it, the discharge is passed through the occlusion tube. This drives the gas into the discharge tube and the X-ray tube again starts working. A high potential difference of the order of 10^5 volts is maintained between the cathode and the anti-cathode. Cathode rays *i.e.*, electrons which start normally from the cathode are converged on the target T striking it with a high velocity. X-ray emission starts from T.

The working of the X-ray tube depends on the pressure inside the glass bulb. If the pressure is comparatively high, the discharge takes place even for a low voltage. The X-rays produced is then very feeble and of low penetrating power (soft X-rays). On the other hand, if the pressure inside is low, a high voltage is needed to maintain the discharge. The X-rays produced then are of high penetrating power (hard X-rays).

The gas-filled tube has several disadvantages. The gas pressure falls due to the absorption of the ions on the walls of the tube. As a result the intensity of the X-rays decreases after some time. Moreover, with this type of tube, it is difficult to control the intensity (as well as quality) of the X-rays due to the *voltage* and *pressure* requirements for the discharge.

(ii) The Coolidge tube: To have an X-ray tube which could produce rays of high penetrating power, it was considered necessary to remove all traces of gas from inside the tube. But even extremely high potentials could not produce such rays inside such a tube *i.e.*, with *cold cathode*. Hence in 1916, Coolidge devised an X-ray tube, called the *Coolidge tube*, which made use of a *hot cathode*.

Although a modern X-ray tube (or Coolidge tube) bears no resemblance to the discharge tube of Roentgen's apparatus, the basic mechanism of X-ray production remains the same. X-rays are still produced by causing cathode rays to strike a solid target, but the techniques have been greatly refined. The essential parts of a modern Coolidge X-ray tube which is widely used for commercial and medical purposes are shown in Fig. 13.2. As there is no need for residual gas to be ionized, the tube is exhausted to the best possible vacuum of the order of 10^{-5} mm of mercury. The source of cathode-ray electrons is a tungsten filament, F which is heated to incandescence either by a storage battery or by a low-voltage alternating current from a step-down transformer, T_2. The filament lies inside a cylinder S of molybdenum (Wehnet cylinder cathode) kept at a high negative potential with respect to the anode. The effect of the cylinder is not only to speed up electrons but also to focus them on a single spot on the anode target.

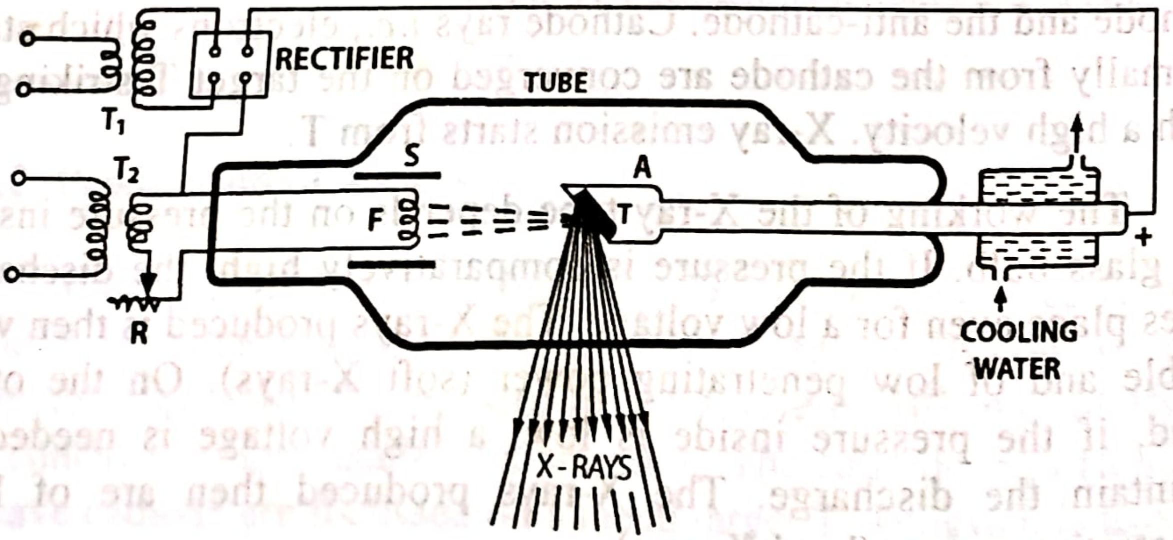


Fig. 13.2

The target T usually employed in X-ray tubes is a massive block of tungsten or in many cases, a molybdenum plug embedded in the face of a solid copper anode. The face of the copper anode is sloped at about 45° to the electron beam. If a high voltage V is applied between the cathode and the anode, with the anode positive, the electrons will be accelerated across the evacuated tube and will acquire very high velocities. For example, electrons accelerated by 100 KV acquire a velocity of 55% of the velocity of light. When these high speed electrons are suddenly stopped by the target, about 99.8 per cent of the energy of the incident beam of electrons is converted into heat and only 0.2 per cent into energy of X-rays. Thus most metals will melt under the terrific bombardment of these high energy electrons. That is why metals like tungsten, platinum and molybdenum, etc. which have high melting points and also have high atomic weight are used as targets. To facilitate the dissipation of enormous amount of heat produced, the usual method is to mount the target on a hollow copper tube through which cold water is continuously circulated. Being very good conductor of heat, copper helps to conduct heat efficiently to the water cooling system.

When the electrons strike the target, X-rays are produced which pass out through a small window in the glass tube as a concentrated beam.

Control of quality and intensity

The *quality* of X-rays is measured in terms of their penetrating power. The penetrating power or quality of the X-ray produced can therefore, be adjusted by varying the potential difference between the anode and the cathode. Greater the accelerating voltage, higher the speed of the striking electrons and consequently, more penetrating the X-rays produced. Highly penetrating X-rays (*i.e.*, those possessing high frequency) are referred to as *hard* X-rays while those having less penetrating power (or low frequency) as *soft* X-rays.

The intensity of X-rays, on the other hand, depends on the number of electrons striking the target. This number is determined by the temperature of the electron-emitting filament which in turn is proportional to the heater current. Hence, by controlling the filament current with the help of a rheostat R, thermionic emission and hence intensity of X-rays can be controlled.

13.3 Origin of X-rays

Roentgen reported in his original paper that X-rays are produced when cathode rays strike some material object. We now know that cathode rays are high-velocity electrons. Therefore, we can restate Roentgen's observation by saying that X-rays are produced when high-velocity electrons strike some material target. Majority of the electrons, however, do nothing spectacular at all. Most of them undergo glancing collisions with the particles of the target material. In this sort of collision, the electrons are repelled by the electron cloud of the atom and therefore deflected and to some extent will be slowed down. The small amount of energy so lost will be transferred to the atom involved, so that by this process the target material gains energy; it is heated up. In course of these collisions, the electrons lose their energy a little at a time and thus merely increases the average kinetic energy of the particles of the target material. The result is that the temperature of the target material is increased. It is found that almost 99.8 per cent of the energy of the electron beam goes into heating the target.

But some of the bombarding electrons produce X-rays by losing their energy in the following two ways:

(i) Some of the bombarding electrons penetrate the interior of the atoms of the target material and are attracted by the positive charge of their nuclei. As an electron passes close to the positive nucleus, it is deflected from its path due to this attractive force as shown in Fig. 13.3. Thus the electron suffers a considerable change in

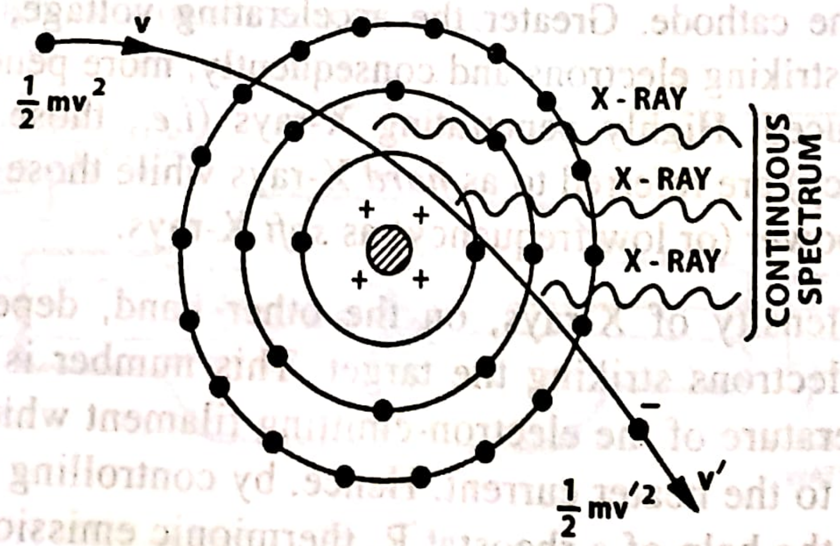


Fig. 13.3

direction and at the same time a large reduction in its speed and therefore energy. The electron therefore, experiences deceleration during its deflection in the strong field of the nucleus or the electron is said to have suffered a *collision.* Now radiation results when a charged body is accelerated (or decelerated). Thus the energy lost during this deceleration is given off in the form of X-rays. How much energy is emitted depends not only on the details of the *collision* but also on how much energy the electron retains after any previous collision. A wide range of photon energies of continuously varying wavelength (and hence frequency) will thus be produced. These X-rays constitute the *continuous spectrum.* The continuous spectrum has a sharply defined short wave-length limit λ_{min} (or high frequency limit ν_{max}). This corresponds to the bombarding electrons which make solid hits and lose most or all of their energy in just one collision. These electrons are rapidly decelerated, resulting in the production of an energetic pulse of electromagnetic radiation. This may be regarded as an *inverse photo-electric effect* in which an electron produces a photon. According to classical electromagnetic theory, there is no lower limit to the wavelength of the radiation that a high velocity electron can produce when it is stopped suddenly. But there is a quantum limit. In chapter 12 we found that photons of given frequency *i.e.*, energy produce electrons with a certain maximum energy. Here we find the reverse phenomenon – electrons of a given energy produce X-ray photons with a certain maximum energy. Both photoelectric effect and X-ray production confirm the quantum view of radiation.

Let the velocity of the electron of Fig. 13.3 be reduced from v to v' during its passage through the atom of the target material. Then the energy lost by the electron is = $(\frac{1}{2}mv^2 - \frac{1}{2}mv'^2)$. This must be equal to the energy of the X-ray photons emitted.

$$\therefore \quad \frac{1}{2} m (v^2 - v'^2) = h\nu$$

The highest or maximum frequency of the emitted X-rays corresponds to the case when the electron is *completely stopped* in just one collision *i.e.*, when v' = 0. In that case

$$\frac{1}{2}mv^2 = h\nu_{max} \qquad \text{(i)}$$

It was shown by Duane and Hunt that if the electron loses all its energy in a single encounter, then the wavelength of the resulting radiation in Angstroms can be calculated in the following manner.

If the electron is accelerated through a potential of V volts, then

$$\frac{1}{2}mv^2 = eV \qquad \text{(ii)}$$

From (i) and (ii), we get

$$h\nu_{max} = eV, \text{ or } \nu_{max} = \frac{eV}{h}$$

$$\text{Now } h\nu_{max} = h\frac{c}{\lambda_{min}}$$

$$\therefore \frac{hc}{\lambda_{min}} = eV; \text{ or, } \lambda_{min} = \frac{hc}{eV}$$

Substituting the values of

$e = 1.602 \times 10^{-19}$ Coulomb

$h = 6.62 \times 10^{-34}$ J-s

and $c = 3 \times 10^8$ m/sec.

we get

$$\lambda_{min} = \frac{6.62\times10^{-34}\times3\times10^8}{1.602\times10^{-19}\,V}$$

$$= \frac{1.24\times10^{-6}}{V}\,m$$

$$= \frac{12,400}{V}\,\text{Å} \;(1\,\text{Å} = 10^{-10} m) \qquad (13.1)$$

This expression gives the minimum wavelength, since no electron can lose more energy than it has. But there will be a

continuous distribution of radiation toward longer wavelengths as there are all sorts of collisions, from direct hits to glancing ones. Thus glancing collisions account for the continuous spectrum of X-rays from any target material and also for the inefficiency of the conversion of the electron energy into X-ray energy. The Germans aptly called this continuous radiation *bremsstrahlung*. The term, which literally means *braking radiation*, is a highly descriptive one, since it refers to radiations that are due to *braking* or *slowing down* of high velocity electrons. Some continuous spectra are shown in Fig. 13.4. These X-rays are independent of the nature of the target material but are determined by the potential difference between the cathode and the anode of the X-ray tube.

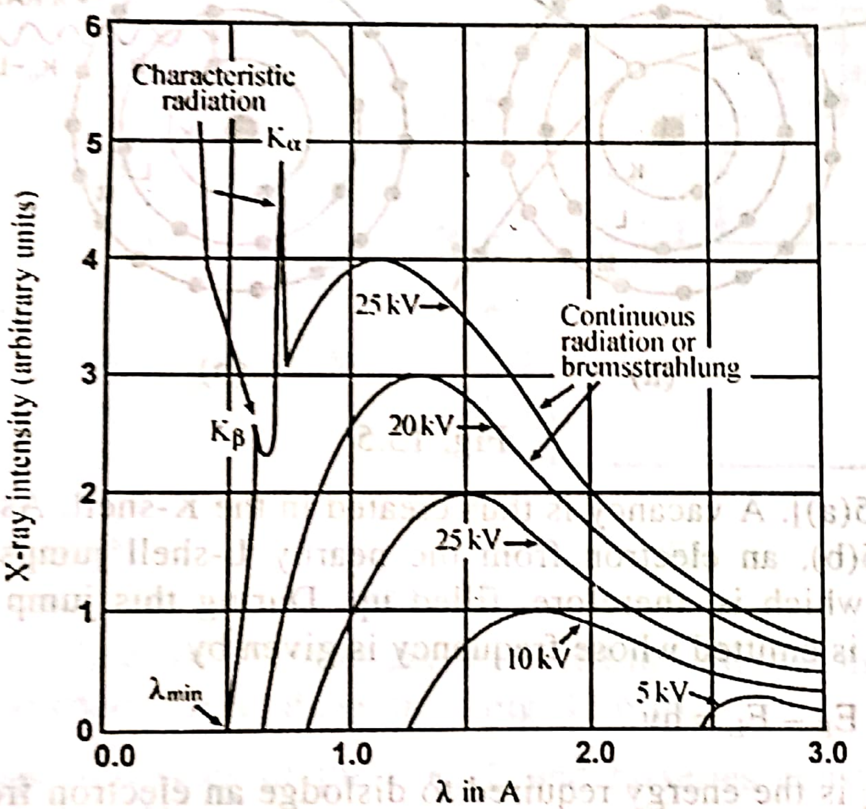


Fig. 13.4

(ii) Looking at the collision process more closely, it can be seen that there is another very important kind of collision energy exchange. Some of the high-velocity electrons while penetrating the interior of the atoms of the target material, knock off the tightly-bound electrons in the innermost shells (like K, L – shells etc.) of the atoms. Although the energy required to ionize an atom by removing an electron is much less than 100 eV, the energy required to ionize by removing an inner electron may be as high as 120,000

eV. X-ray producing electrons usually have enough energy to produce ions by removing inner electrons from the atom, even down to the innermost or K-shell. When electrons from outer orbits jump to fill up the vacancy so produced, the energy difference between the two orbits is given off in the form of X-rays of *definite* wavelength (and frequency). These X-rays constitute the line spectrum which is *characteristic of the material of the target.*

Fig. 13.5 illustrates the origin of characteristic X-rays. The high velocity incident electron knocks off one electron from the K-shell

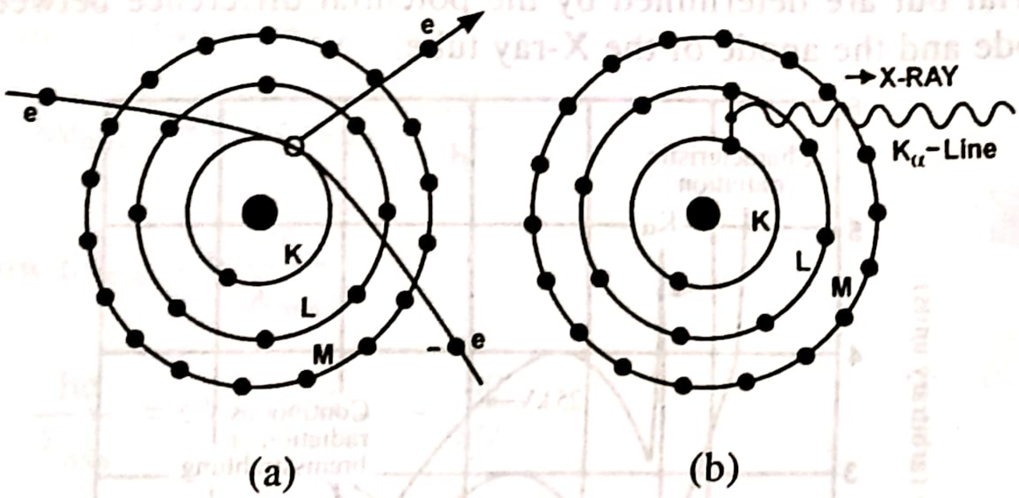


Fig. 13.5

[Fig. 13.5(a)]. A vacancy is thus created in the K-shell. As shown in Fig. 13.5(b), an electron from the nearby L-shell jumps into this vacancy which is, therefore, filled up. During this jump an X-ray radiation is emitted whose frequency is given by

$$E_K - E_L = h\nu$$

where E_K is the energy required to dislodge an electron from the K-shell and E_L is that required for L-shell. Since the energy difference between the K-shell and L-shell is very large, the X-rays emitted possess very high energy and is therefore very penetrating. This is referred to as K_α line.

If the vacancy in the K-shell is filled up by an electron from M-shell, instead of L-shall, the X-ray emitted would be still more energetic and hence more penetrating since the energy difference (E_K – K_M) is more than the energy difference (E_K – E_L). This is referred to as K_β line.

Although there is possibility that the vacancy in the K-shell may be filled up by electrons in still higher orbits, usually two lines – K_α and K_β, out of many lines of this series are detected.

Similarly, when the incident electron carries somewhat less energy, it will dislodge an electron from an L-shell. An electron from either M-shell or other outer orbits will then fill up this vacancy giving rise to X-rays of frequency lower than those of the K-series. These X-rays constitute another series, called the L-series of the X-ray spectrum as shown by L_α, L_β and L_γ lines in Fig. 13.6(a).

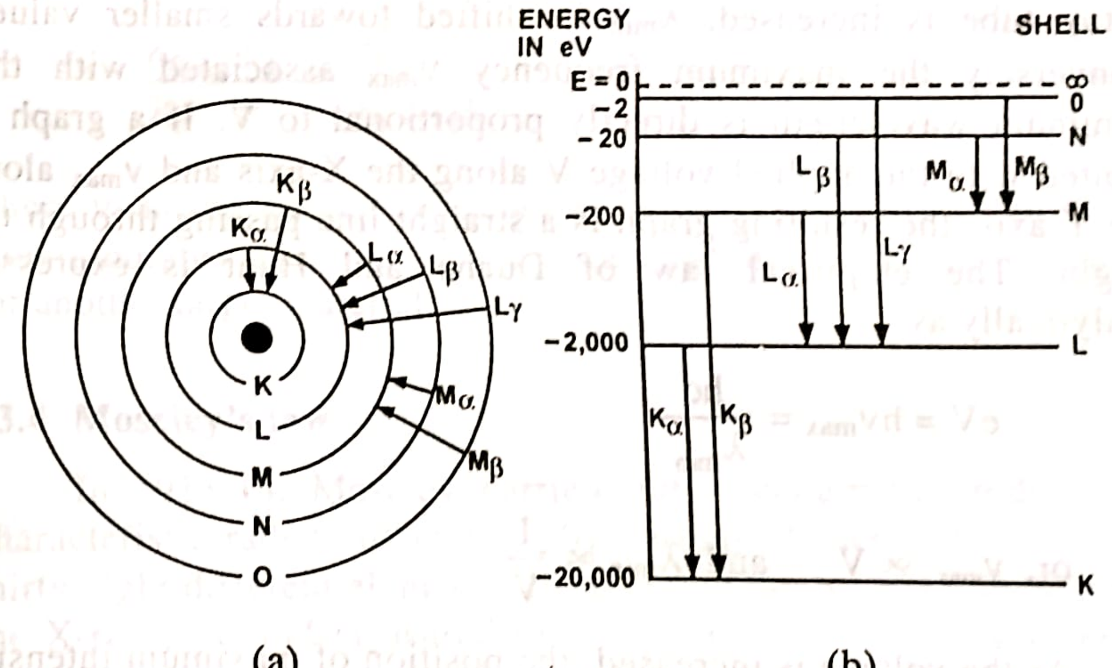


Fig. 13.6

Spectral lines of M-series are produced in a similar manner as shown in the energy-level diagram of Fig. 13.6(b).

The spectral lines of these K, L and M-series constitute the line spectra of the X-rays and *are characteristic of the material* used as target in the X-ray tube. They are also referred to as *characteristic X-rays* of the target material.

Thus the X-rays produced by an X-ray tube consist of two parts:

(i) Continuous spectrum

One part consists of a series of uninterrupted wavelengths produced due to the deceleration of the high-velocity electrons when they are deflected while passing near the positively charged nucleus of an atom of the target material. This is the **continuous spectrum**.

It has a sharply-defined short-wavelength limit (λ_{min}) below which no radiation is emitted. Above this limit, the intensity of the radiation increases rapidly with increasing wavelength and after reaching a maximum, decreases gradually. The intensity never reaches zero showing that the radiation contains all possible wavelengths above the minimum limit.

Duane and Hunt showed that the cut-off wavelength λ_{min} is inversely proportional to the potential difference V between the cathode and the anode. In other words, when the voltage across the X-ray tube is increased, λ_{min} is shifted towards smaller values. Conversely the maximum frequency ν_{max} associated with this minimum wavelength is directly proportional to V. If a graph is plotted with the applied voltage V along the X-axis and ν_{max} along the Y-axis, the resulting graph is a straight line passing through the origin. The empirical law of Duane and Hunt is expressed analytically as

$$eV = h\nu_{max} = \frac{hc}{\lambda_{min}}$$

$$\text{or,}\quad \nu_{max} \propto V \quad \text{and}\quad \lambda_{min} \propto \frac{1}{V}$$

As the voltage is increased, the position of maximum intensity is shifted towards short wavelength side. The intensity of the continuous spectrum, given by the area enclosed by the curve of Fig. 13.6, is found to be very nearly proportional to the *square of the applied voltage* for a given target and to the atomic number of the target material when the applied voltage is kept constant.

(ii) Line spectrum

Line spectrum is produced when the incident high energy electron knocks out an electron from the innermost orbit of the target atom followed by electron jump from higher orbits to fill up the vacancy. It consists of discrete spectral lines which constitute K-series. L-series, M-series, etc. Out of these, spectral lines constituting the K-series are the most energetic hence most penetrating. These form the *hard* X-rays. L and M series are less penetrating and form the *soft* X-rays.

Line spectrum is the *characteristic* of *the target material used.* The *number* of lines present in the spectrum depends both *on the nature of the target material and the excitation voltage.*

For each target these is a minimum potential below which the line spectra do not appear. This critical potential difference below which no line spectra appear, is different for different targets. The K-spectrum is shifted towards the shorter wavelength side as the atomic number of the target is increased. The exact relationship, as found by Mosley, is

$$\frac{\lambda_2}{\lambda_1} = \frac{\nu_1}{\nu_2} = \frac{(Z_1 - 1)^2}{(Z_2 - 1)^2}$$

where ν_1 is the frequency of the K_α – line for a target material of atomic number Z_1 while ν_2 and Z_2 are the corresponding quantities for another target material.

13.4 Moseley's law

In 1913-14, Moseley carried out a systematic study of the characteristic radiations emitted by various elements. He employed thirty eight different elements, from aluminium to gold, as targets in the X-ray tube and measured the frequencies of the X-rays emitted from these elements. Moseley found that, unlike optical spectra, the X-ray characteristic spectra of different elements are remarkably similar to each other in the sense that each consists of K-, L- and M-series. There is, however, one very important difference. The frequency of lines (in every series) produced from a heavier element is greater than that produced from a lighter element. In order to establish a possible relationship between the frequency of a given spectral line (ν) and some property of the atom emitting it, Moseley first plotted the square root of the frequency ($\sqrt{\nu}$) of a spectral line with atomic weight (A). He observed that the frequency did not vary uniformly with atomic weight. He then plotted the square root of the frequencies ($\sqrt{\nu}$) of a given spectral line, say K_α, against the atomic numbers (Z) of the elements emitting that line. A straight line as shown in Fig. 13.7 was obtained. The same linear relationship was found to hold good for any line in any series. Moseley therefore

concluded that it is the *atomic number* and not the atomic weight which is the fundamental property of the elements.

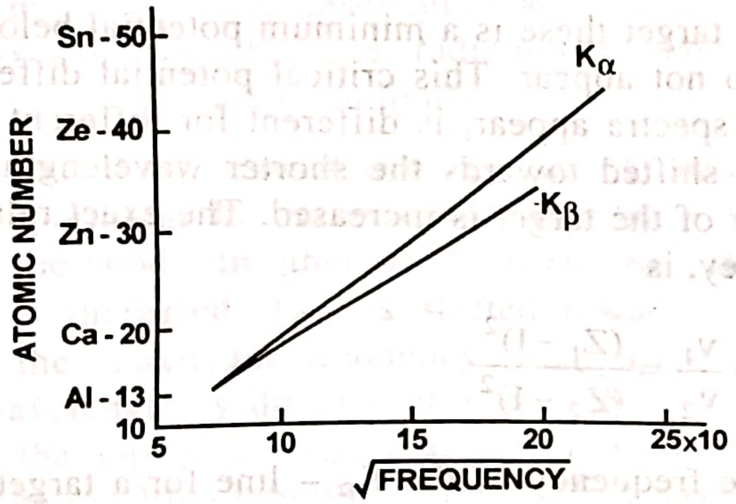


Fig. 13.7

Statement of Moseley's law :

From the conclusion drawn above Moseley stated that *the frequency of a spectral line* in the characteristic X-ray spectrum, varies directly as the square *of the atomic number of the element emitting it.*

Mathematically,

$$\sqrt{\nu} \propto Z \qquad \text{or, } \nu = a Z^2 \qquad (13.2)$$

where a is a constant for a particular series but varies from one series to another, *i.e.*, their values for K-series are different from those for L-series, etc.

Moseley was a contemporary of Niels Bohr. Both were working in Rutherford's laboratory in Manchester, England. By extending Bohr's theory of the hydrogen atom to heavier elements, Moseley showed that the law as given by eqn. 13.2, is a transition from n = 2 to n = 1 giving rise to the K_α radiation as explained earlier. Moseley's law can be arrived at as follows:

The energy of an electron in the n^{th} orbit as given by Bohr's theory is

$$W_n = -\frac{RchZ^2}{n^2} \tag{13.3}$$

When the electron jumps from the n_i to the n_f energy state, the frequency of the emitted line is given by

$$h\nu = W_i - W_f = RchZ^2\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

For the simplest case of electron jump from L-shell to the K-shell *i.e.*, from the energy states corresponding to $n_i = 2$ to that corresponding to $n_f = 1$, we get

$$\nu = RcZ^2\left(\frac{1}{1} - \frac{1}{4}\right)$$

$$= \frac{3}{4} RcZ^2$$

where R is the Rydberg constant.

$$\therefore \quad \nu \propto Z^2$$

$$\text{or,} \quad \sqrt{\nu} = aZ$$

which is Moseley's law.

In arriving at Moseley's law, Bohr's theory has been applied to the inner filled K - and L - shells whereas it is actually applicable to hydrogen or hydrogen like atoms. A correction is therefore warranted.

According to Moseley, the one electron remaining in the K-shell during the transition of the hole from the K to L shells shields the nucleus so that the effective Z is reduced to (Z – 1) for the emission of the K_α line. Similarly, the nine electrons remaining in the K and L shells shield the nucleus during the transition of a hole from the L to M shells. However, this shielding is not perfect and the effective Z for the emission of L_α line is reduced only to (Z - 7.4).

Thus, the presence of electrons in an inner shell screens the effect of the nucleus on the electrons of the outer shell. This screening is taken into account by replacing the charge of the

nucleus Ze by (Z – σ)e while applying Bohr's theory to the study of the characteristic X-rays. σ is a constant called the *screening parameter* whose value depends upon the particular line *i.e.*, the shell. When this is taken into account Moseley's law becomes

$$\sqrt{\nu} = a\,(Z - \sigma)$$

The frequencies of the various lines in the K- and L- series of the characteristic X-ray spectrum are given by the following empirical formulae.

K - series

The general formula is

$$\nu = Rc\,(Z-1)^2\left(1 - \frac{1}{n^2}\right) \text{ where } n = 2, 3, \text{ etc.}$$

Here the nuclear screening constant (σ) is unity.

(i) For K_α line

$$\nu = Rc\,(Z-1)^2\left(1 - \frac{1}{4}\right) = \frac{3}{4}\,Rc\,(Z-1)^2$$

(ii) For K_β line

$$\nu = Rc\,(Z-1)^2\left(1 - \frac{1}{9}\right) = \frac{8}{9}\,Rc\,(Z-1)^2$$

L - series

The general formula is

$$\nu = Rc\,(Z-7.4)^2\left(\frac{1}{4} - \frac{1}{n^2}\right) \quad n = 3, 4, \text{ etc.}$$

Here, the nuclear screening constant is 7.4

(i) For H_α line, n = 3

$$\therefore\ \nu = Rc\,(Z-7.4)^2\left(\frac{1}{4} - \frac{1}{9}\right)$$

$$= \frac{5Rc}{36}\,(Z-7.4)^2$$

(ii) For H_β line, n = 4

$$\therefore \quad \nu = Rc\,(Z - 7.4)^2 \left(\frac{1}{4} - \frac{1}{16}\right)$$

$$= \frac{3Rc}{16}\,(Z - 7.4)^2$$

Importance of Moseley's law

The original ordering of the elements in the periodic table was on the basis of their atomic weights. But Moseley's law proved for the first time that it is the atomic *number* and not the atomic *weight* of an element which determines its characteristic properties (both physical and chemical). When the wavelengths of the K_α lines were arranged in the order of atomic weights, Moseley found that the sequence was not perfect since there were both gaps and wavelengths out of order. He attributed the gaps to undiscovered elements and proposed that there should be a unique correlation between the wavelength series and atomic number. This led Moseley to surmise that there is in the atom a fundamental quantity, which increases by regular steps as we pass from one element to the next. This quantity can only be the charge *on the central positive nucleus*. Moseley's finding provides the proper guideline that elements must be arranged in the periodic table according to their atomic *numbers* and not their atomic *weights* as suggested earlier by Mendeleeff. In Moseley's days, nickel with an atomic weight of 58.69 was listed ahead of cobalt with an atomic weight of 58.94 although there was some chemical evidence that the order of these two elements should be reversed. This was demonstrated by Moseley who showed that the atomic number of cobalt is 27 while that of nickel is 28. The anomaly regarding the relative positions of potassium ($_{19}K^{39}$) and argon ($_{18}A^{40}$) was similarly resolved.

Moseley's work has also helped to perfect the periodic table. The missing elements according to Moseley were those with atomic numbers of 43, 61 72 and 75. Element 72, hafnium was discovered in 1925. The two final elements, numbers 43 and 61 have no stable isotopes, but were produced artificially. Technetium (43) was discovered in 1937 and has an isotope (A = 97) with a half-life of 2.6×10^6 years. Promethium (61), which has an isotope (A = 145) with a half-life of 18 years, was discovered in 1947.

13.5 Measurement of X-rays

The methods mentioned below can be used for measuring the intensity of X-rays.

(i) *By measuring fluorescent light:* Roentgen first detected X-rays by the fluorescence they produce in certain materials. This effect can be used for quantitative measurement of the intensity of X-rays if it is coupled with an objective measurement of the fluorescent light produced.

(ii) *By using a photographic plate* : Roentgen also observed that X-rays blacken a photographic plate when incident on it, and this may be used to measure the intensity of X-rays. The degree of blackening produced on a photographic film due to X-ray exposure can be measured by a densitometer and this can be correlated to the intensity of the incident X-rays.

(iii)*By measuring ionization current* : This method was suggested by Roentgen's experiments on the conductivity of air caused by X-rays. X-rays ionize a gas while passing through it. This effect can be measured quantitatively with an ionization chamber. Stronger the ionization current the greater the intensity. There are many forms of ionization chambers one of which is shown schematically in Fig. 13.8. The chamber itself is a metallic box into which an electrode AB is inserted through an insulating plug. The electrode is usually maintained at a positive potential, so that any free

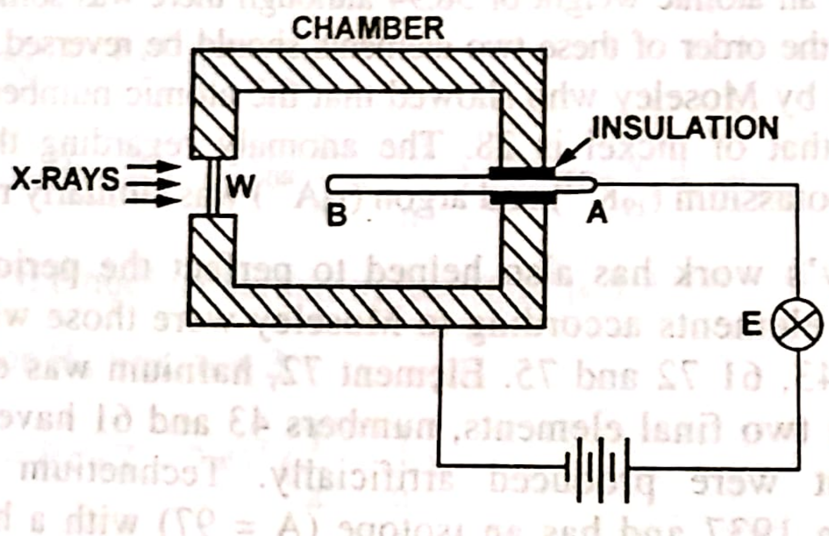


Fig. 13.8

electrons within the chamber will be attracted to it. These electrons are produced when X-rays are allowed to enter the chamber through the window. The potential difference should be high enough so that charges are collected before the electrons recombine with positive ions and at the same time low enough so that the dielectric strength of the air prevents a discharge. When collected by the electrode AB, the electrons constitute a small electric current which is proportional to the intensity of the incident X-rays. This current can be measured by a sensitive electrometer which quickly responds to any changes in the X-ray intensity. In modern instruments, the current is amplified electrically and displayed by a microammeter which can be directly calibrated in terms of X-ray intensity.

13.6 Absorption of X-rays

The most spectacular property of X-rays is their ability to penetrate materials that are opaque to less energetic radiation. However, all materials through which X-rays pass absorb them to some extent. When a narrow and monochromatic beam of X-rays passes through a material, part of it is absorbed and the remaining part is transmitted. Absorption of X-rays can be studied with the help of an ionization chamber. A schematic diagram of the apparatus is shown in Fig. 13.9. X-rays produced by an X-ray tube are collimated into a well-defined narrow beam by passing them through two fine

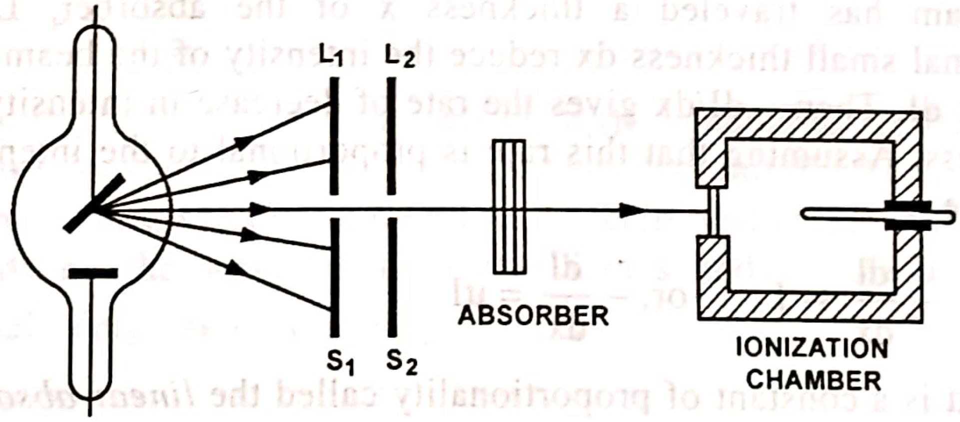


Fig. 13.9

slits S_1 and S_2 in two lead plates. The beam is then rendered monochromatic by Bragg reflection from a crystal, and allowed to enter the ionization chamber. The ionization current is measured which gives the intensity of the X-rays.

Next, a sheet of the absorbing material of given thickness is interposed in the path of the X-ray beam before it enters the ionization chamber. The ionization current and hence the intensity of the X-ray beam is measured which is found to be reduced. The thickness of the sheet of the absorbing material is increased in some given steps and the intensity corresponding to each thickness is measured. When the transmitted intensity is plotted against thickness, a graph as that shown in Fig. 13.10 is obtained. This is an exponential decay curve which can be derived in the manner described below.

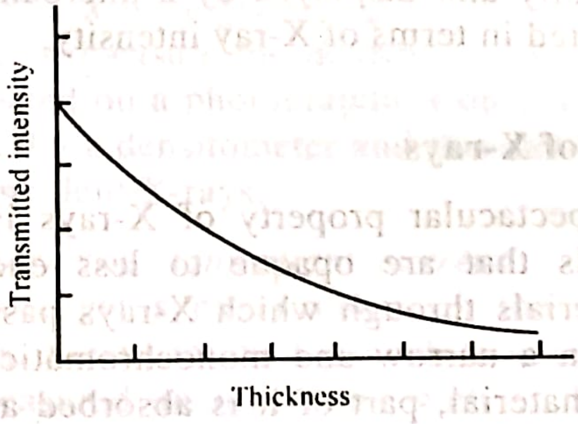


Fig. 13.10

Let I_0 be the intensity of the X-ray beam incident normally on an absorber sheet and I the intensity of the transmitted beam after the beam has traveled a thickness x of the absorber. Let an additional small thickness dx reduce the intensity of the beam by an amount dI. Then – dI/dx gives the rate of decrease in intensity with thickness. Assuming that this rate is proportional to the intensity I, we have

$$-\frac{dI}{dx} \propto I; \quad \text{or,} \; -\frac{dI}{dx} = \mu I$$

where μ is a constant of proportionality called the *linear absorption coefficient* of the absorber (also known as the *macroscopic absorption coefficient and linear attenuation coefficient*)

The above relation can be written as

$$\frac{dI}{I} = -\mu dx \qquad (13.4)$$

Integrating both sides of the above equation, we get

$$\int \frac{dI}{I} = -\mu \int dx\,; \qquad \text{or,} \quad \log_e I = -\mu x + C \tag{13.5}$$

where C is the constant of integration. The value of C can be found from the initial conditions which are that when $x = 0$, $I = I_o$. Substituting these values in eqn. 13.5 we have

$$C = \log_e I_o$$

Hence eqn. 13.5 becomes

$$\log_e I = -\mu x + \log_e I_o$$

$$\text{or,} \quad \log_e \frac{I}{I_o} = -\mu x$$

$$\text{or,} \quad \frac{I}{I_o} = e^{-\mu x}$$

$$\text{or,} \quad I = I_e e^{-\mu x} \tag{13.6}$$

Thus the intensity of the X-ray beam decreases exponentially with the thickness of the absorbing material as shown in Fig. 13.10.

By rearranging eqn. 13.4 above, we get

$$\mu = \frac{-dI/I}{dx} \tag{13.7}$$

Hence, the linear absorption coefficient (μ) is equal to the fractional decrease in the intensity of the X-rays per unit thickness of the absorber. Thus it has the dimensions of reciprocal length. Value of μ depends on the wavelength of the X-rays and on the absorbing material. Very nearly, $\mu \propto \lambda^3$.

13.7 Half-value Layer

One consequence of the exponential attenuation is that it is impossible to reduce an X-ray beam to nothing – no matter how thick the material inserted into the beam is. There will always be a chance, albeit very small, of some radiation being transmitted. As

can be seen from the shape of the attenuation curve of Fig. 13.10, it is also not possible to state a thickness for the *range* of the radiation. Nevertheless, the penetrating power or *quality* can be defined, quite conveniently, in terms of a thickness. This thickness is the so-called half-value layer (H.V.L), sometimes called the half-value thickness. It is the thickness of a given material which will reduce a narrow beam of X-rays to one-half of its original value. Let this thickness be $D_{½}$. Then, replacing x by $D_{½}$ in eqn. 13.6 we have

$$I = I_o e^{-\mu D_{1/2}}; \qquad \text{or, } I/I_o = e^{-\mu D_{1/2}}$$

But I/I_o by definition is ½. then we get

$$\frac{1}{2} = e^{-\mu D_{1/2}}; \qquad \text{or, } e^{\mu D_{1/2}} = 2$$

$$\text{or, } \mu D_{1/2} = \log_e 2$$

$$\text{or, } \mu = \frac{\log_e 2}{D_{1/2}} = \frac{0.693}{D_{1/2}} \qquad (13.8)$$

Various thicknesses of the test material are inserted into the narrow beam and an attenuation curve like that shown in Fig. 13.10 is obtained. From this curve the H.V.L. can be read off in the same way as the thickness for any other fractional reductions. The linear attenuation coefficient can then be determined with the help of eqn. 13.8.

If a thickness $D_{½}$ reduces the beam to a half, then an addition of equal thickness, making $2D_{½}$ in all, will reduce this half by a further half, and the whole beam to quarter of its original value. This thickness $2D_{½}$ is called the quarter-value layer (Q.V.L.) and is twice the thickness of the H.V.L. – at least for a photon beam of one energy. By the same token, a thickness of $3D_{½}$ reduces the beam to one-eighth, and $4D_{½}$ brings it down to one-sixteenth and so on. Mathematically it can be shown that a thickness of $n \times D_{½}$ will reduce the beam to $(½)^n$ of its original value. if n = 10, the beam is reduced to $(\frac{1}{2})^{10}$ or roughly to 1/1000 of its original value.

The absorption coefficient (μ) depends strongly on the density of the absorbing material which, of course, changes greatly if the material goes from gaseous to solid states. We can write

$$\mu x = \frac{\mu}{\rho} \times x\rho$$

The quantity μ/ρ or μ_m which is called the *mass absorption coefficient* has been found to be more important for practical application. Its units are of the form of area divided by mass (cm^2/gm). It may be defined as equal to the fractional decrease in absorbing material. Dimensional analysis shows that $(x.\rho)$ is mass per unit area m_a. It is the mass of a sheet or slab of the absorber which has a thickness x and a unit of surface area normal to the incident X-ray beam. In these terms μx becomes $\mu_m m_a$. Eqn. 13.6 becomes

$$I = I_o e^{-\mu_m m_a} \tag{13.9}$$

μ_m has also been, found to depend strongly on the wavelength of X-rays and the atomic number of the absorber material, the exact relationship being

$$\mu_m = k\lambda^3 Z^3$$

13.8 Wave nature of X-rays

The physical nature of X-rays remained in doubt for many years after their discovery by Roentgen in 1895. Some thought X-rays to be high speed particles like cathode rays but more penetrating while others regarded them to be electromagnetic waves of extremely short wavelength. Many experiments were performed to test these two hypotheses but without any success. X-rays could not be deflected in electric and magnetic fields thereby showing that they are not charged particles like cathode rays neither they could be reflected or refracted in earlier experiments thereby demonstrating wave nature. To test their wave nature attempts were made to produce interference as well as diffraction effects similar to those observed with light waves. Early experiments on the diffraction of X-rays were performed by Haga and Wind, and Walter and Phol. The results of Walter and Phol on the diffraction of X-rays by a wedge-shaped slit, were employed by Sommerfeld to calculate the wavelength of X-rays which he found to lie between 10^{-8} cm and 10^{-9} cm. These wavelengths were considerably shorter than that of light

waves. Hence, it was soon realized that to produce the same order of diffraction as that produced by light waves with a transmission grating, the X-ray grating would require 40 million rulings per cm! While these experiments on X-ray diffraction were still in the development stage, it occurred to German theoretical physicist Max Von Laue in 1912 that a natural crystal can be used as a closely-spaced three-dimensional grating. At Laue's suggestion, the crucial experiment of using a crystal as a diffraction grating was performed by W. Friedrich and P. Knipping. Their experimental arrangement is shown in Fig. 13.11. The X-ray beam from a Coolidge tube is collimated into a fine pencil of rays by a pinhole in each of the two lead screens L_1 and L_2. The beam is then allowed to be incident on a thin crystal of zinc sulphide (ZnS) along one of its axes of symmetry. After passing through the crystal the X-ray beam is allowed to fall on a photographic film placed beyond the crystal.

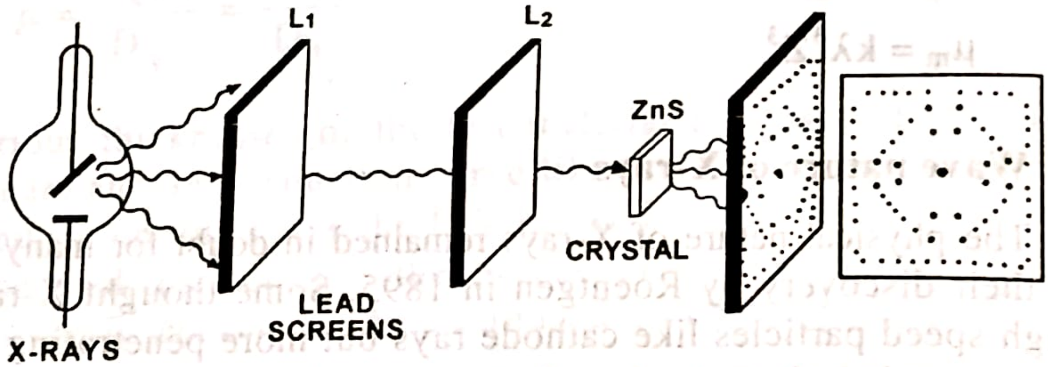


Fig. 13.11

Most of the X-rays go directly through the crystal and produce a black spot at the centre of the film. This central dark spot is surrounded by a group of symmetrically situated dark sports which are due to the many other weak diffracted X-ray beams emerging in different directions. This symmetrical pattern of spots is known as Laue pattern and the spots are called Laue spots. The positions of these spots were found to change with the change in the orientation of the crystal and were different for different crystals. The wavelength of the heterogeneous X-ray beam used in this

experiment was calculated to be from 0.13 to 0.48 Å. *This experiment proved the wave-nature of X-rays* and also confirmed the supposition that crystals have their atoms arranged in a regular structure.

Laue patterns are widely used for the following purposes:

(i) in locating as well as determining angular relationship between different planes in a crystal lattice which are rich in atoms.

(ii) in determining the microscopic structure of matter in solid state, as for example in metallurgy, for studying the mechanical processes of rolling, hardening and annealing.

13.9 Bragg's law

Late in 1912, shortly after the Laue experiment, William L. Bragg along with his father William H. Bragg devised another technique for diffracting X-rays. Instead of observing the effect created by passing the rays through a crystal, Bragg considered how X-rays are scattered by the atoms in the crystal plane. Let an X-ray wave front be incident on a surface row of atoms in a crystal plane as shown in Fig. 13.12. Each atom becomes a source of scattered X-radiation. As the scattered X-rays from all the atoms in the crystal

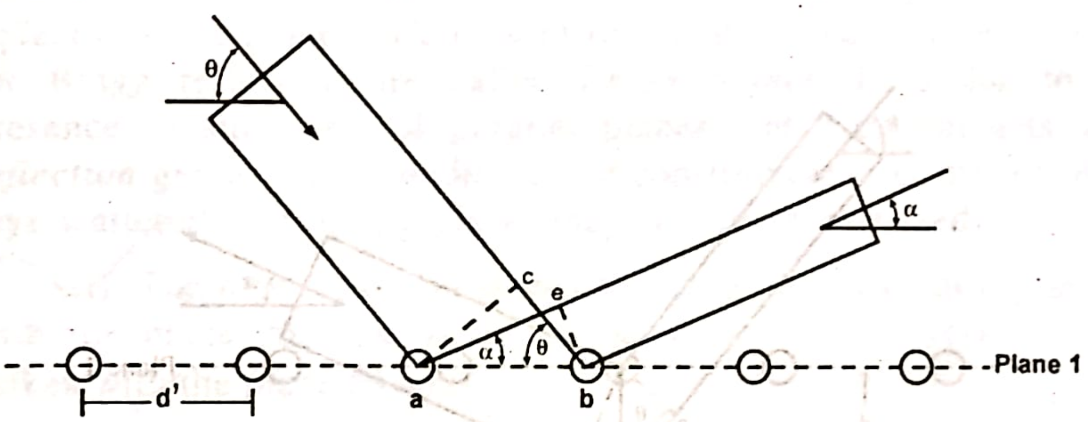


Fig. 13.12

fall on top of one another in a random manner, they generally combine destructively. However, if certain conditions are met, constructive interference will occur at a few places. One of the relations that must be satisfied for constructive interference can be derived as follows.

Let d' be the distance between adjacent atoms and θ and α be the angles of incidence and reflection respectively. As can be seen from the figure, unlike in optics where the angles of incidence and reflection are measured between the rays and the normal to the surface, the Bragg angles θ and α are measured between the incident rays and the surface plane and between the scattered rays and the surface plane respectively. To obtain path difference between rays from adjacent atoms, let us drop perpendiculars ae and bc on the incident and scattered rays respectively. Obviously, this path difference is then ae – cb, and, for reinforcement, *i.e.*, constructive interference, this difference must be some integral multiple of the wavelength of the X-ray. Therefore, for constructive interference, we have

$$ae - cb = m\lambda$$

$$\text{or,}\quad d'\cos\alpha - d'\cos\theta = m\lambda \qquad (13.10)$$

The other relation which must be satisfied for maximum reinforcement is that the scattered rays from successive planes of atoms meet in phase. Referring to Fig. 13.13, in which d is the vertical distance between successive planes, it may be noted that rays scattered from the second plane travel a greater distance than those from the first plane. In order that the rays from successive planes

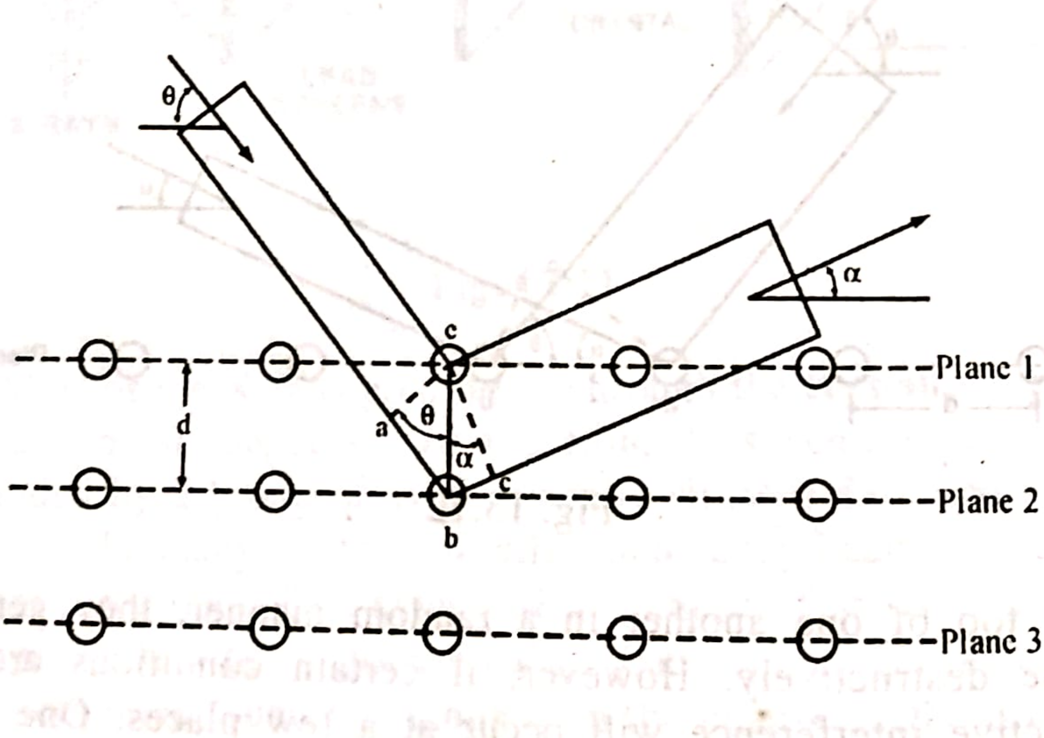


Fig. 13.13

reinforce one another, it is necessary that these additional distances *i.e.*, the path difference, must be some integral multiple of X-ray wavelength. If ea and ec be dropped perpendiculars to the incident and the scattered rays respectively, then the difference in path between a ray scattered from the second plane and a ray scattered from the first plane is given by the sum of the distances ab and dc. For constructive interference to take place,

$$ab + bc = n\lambda$$

$$\text{or, } d\sin\theta + d\sin\alpha = n\lambda \tag{13.11}$$

In general, the conditions imposed by eqns. 13.10 and 13.11, cannot be satisfied simultaneously without considering the scattered wavelets which are in the various layers but not in the plane of incidence. Both conditions are met, however, in the special case where $\theta = \alpha$. Then eqn. 13.10 reduces to zero and eqn. 13.11 becomes

$$2d \sin\theta = n\lambda \tag{13.12}$$

where n is the order of the spectrum. n = 1, 2, 3, etc. refers to the first-order, second-order, third-order maxima respectively.

The relation given by eqn. 13.12 is known as Bragg's law. When $\theta = \alpha$, we have precisely the condition of regular optical reflection. Because of this, Bragg scattering is usually called Bragg *reflection*. The planes of atoms in the crystal which are responsible for Bragg reflection are called *Bragg planes*. It is due to the presence of such sets of parallel planes that a crystal acts as a *reflection grating*. The conditions for constructive interference of X-rays scattered from Bragg planes may now be summarized.

(i) The *first condition* is that the angle the incident ray makes with the plane must be equal to the angle that the reflected ray makes with the plane ($\theta = \alpha$).

(ii) The *second condition* is that the reflections from successive Bragg planes must meet in phase, *i.e.*, must satisfy the relation, $2d \sin\theta = n\lambda$.

13.10 Bragg's X-ray spectrometer

For studying X-ray diffraction the Braggs devised an X-ray spectrometer in which a crystal is used not as a *transmission grating*

(as in Laue's experiment) but as a *reflection grating*. A schematic diagram of the spectrometer is shown in Fig. 13.14. The X-rays coming from the tube at the left are collimated into a narrow beam by two fine slits S_1 and S_2 cut into two lead plates L_1 and L_2 which absorb all rays

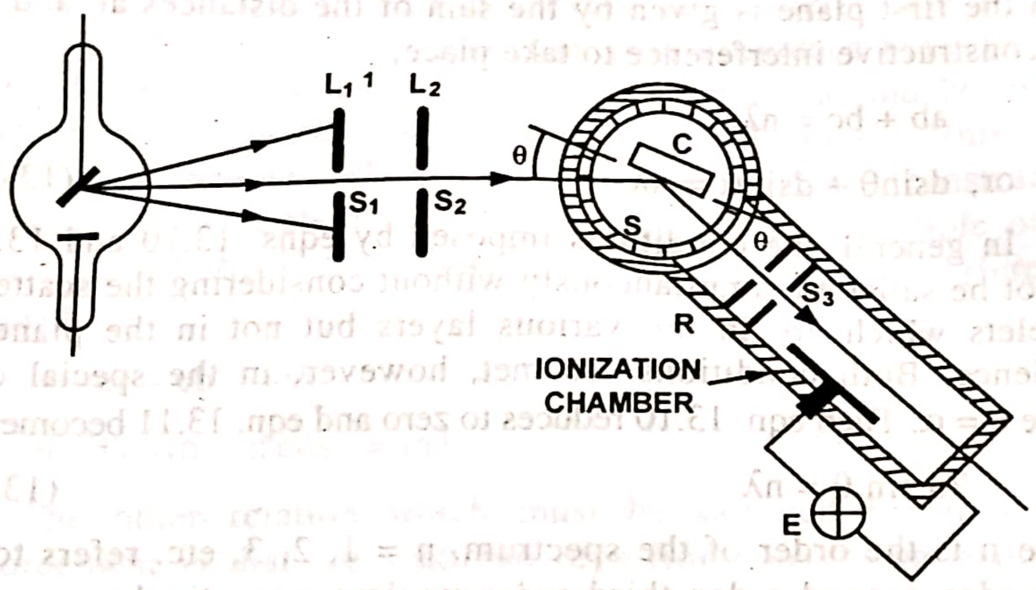


Fig. 13.14

except those which pass through the slits. These rays fall on a crystal C which is mounted on the turntable T of the spectrometer. The turntable and hence the crystal is capable of rotation about an axis parallel to the slit and perpendicular to the plane of the figure. The angle of rotation can be read from the circular scale.

Most of the incident X-rays pass straight through the crystals. However, some X-rays are scattered by the regularly-arranged atoms in different crystal planes. These scattered X-rays can be looked upon as having been reflected from the crystal plane – particularly those which are rich in atoms. The reflected X-ray beam enters into an ionization chamber carried by an arm R which is capable of rotation about the same axis as the turntable.

The ionization current produced by the reflected X-ray beam can be measured by a sensitive electrometer E. When the electrometer readings are plotted against the corresponding angles of incidence θ, curves similar to those in Fig. 13.15 and 13.16 are obtained. Fig. 13.15 shows the spectrum when the X-ray beam is

monochromatic while Fig. 13.16 represents the case when the X-ray beam consists of two wavelengths λ_1 and λ_2.

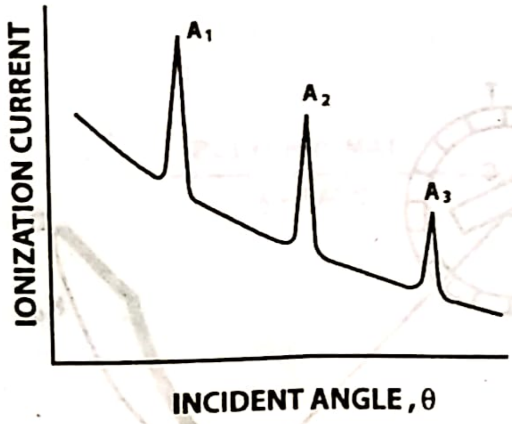


Fig. 13.15

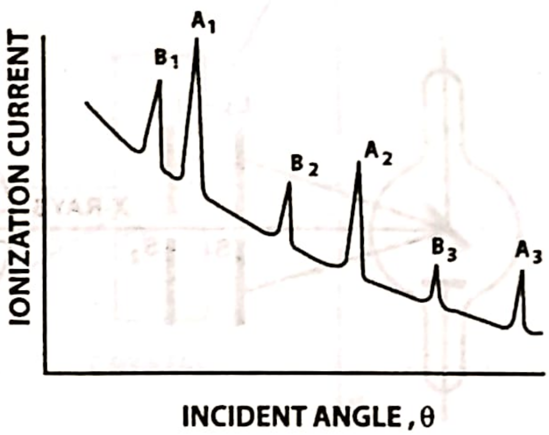


Fig. 13.16

Different directions *i.e.*, angles of incidence θ, for which intense reflections will be produced can be found from Bragg's law.

For the first maxima (n = 1), $\sin\theta_1 = \dfrac{\lambda}{2d}$

For the second maxima (n = 2), $\sin\theta_2 = \dfrac{2\lambda}{2d}$

For the third maxima (n = 3), $\sin\theta_3 = \dfrac{3\lambda}{2d}$

and so on.

As can be seen from Fig. 13.16, as the order of the spectrum increases, the intensities of maxima goes on decreasing. Also, it can be seen that as the order of the spectrum increases, the separation between maxima for λ_1 and λ_2 increases.

Bragg's law can be used to determine the wavelength λ of the X-rays if inter planar spacing d is known. Conversely, d may be computed if the wavelength λ of the X-rays is known from some other experiment. In modern X-ray spectrometers, the ionization

chamber is replaced by a photographic plate set at right angles to the reflected beam as shown in Fig. 13.17. The spectrometer is then referred to as X-ray spectrograph and takes much less time for obtaining the complete spectrum.

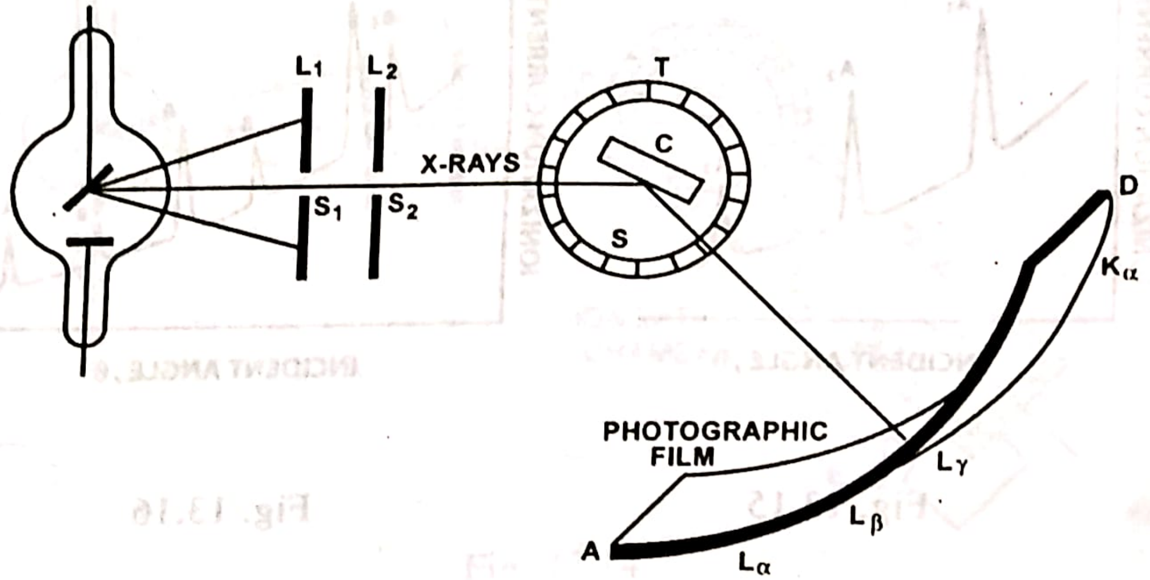


Fig. 13.17

13.11 X-ray crystallography – Laue method

X-ray diffraction in crystals has become a powerful, indeed the most powerful, method for studying the structure of crystals. The Bragg method is better for the study of the wavelengths of X-rays, but the Laue technique is very useful for the study of the crystals.

In Laue method for the study of crystal structure, a bundle of *polychromatic* X-rays impinge at 90° on the plane of the crystal under study (Fig. 13.18). On their way through the crystal, X-rays encounter Bragg's planes with different spacings d. These different families of planes make different angles θ with the direction of the X-rays. Some combinations of d, θ and λ will satisfy the Bragg relation with the resulting increase in the intensity of the diffracted X-rays for this combination.

As can be seen from Fig. 13.18, X-rays which pass straight through the crystal without suffering any diffraction produce the intense black spot at point C on the photographic plate PP. Different

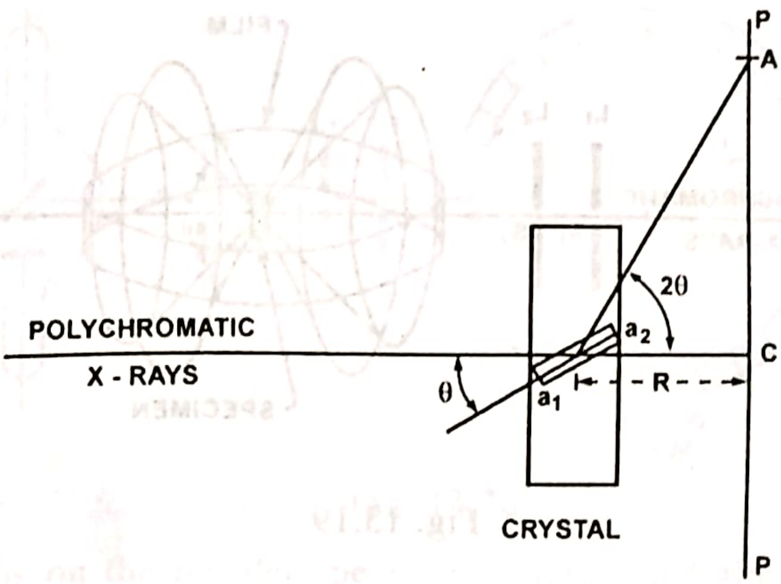


Fig. 13.18

wavelengths present in the primary beam will fall on different Bragg planes at various angles of incidence and those satisfying the Bragg condition will produce somewhat less pronounced spots around this central spot.

Let a_1a_2 represent the position of one of the possible Bragg plane, θ the corresponding angle of incidence and A the spot on the photographic plate. Then from the figure, we have

$$AC = R \tan 2\theta$$

AC and R can be measured. θ can then be found for the corresponding plane.

13.12 Powder crystal method

The Bragg and Laue methods require single crystals large enough for study. There is another X-ray crystallographic technique which is used when *single crystals of reasonably large size are not available*. This technique, known as the Debye – Scherrer method, requires that the crystalline substance be finely ground and powdered so that its tiny crystals assume random orientations. A small specimen of this crystalline powder is interposed in the path of a narrow monochromatic beam of X-rays as shown in Fig. 13.19.

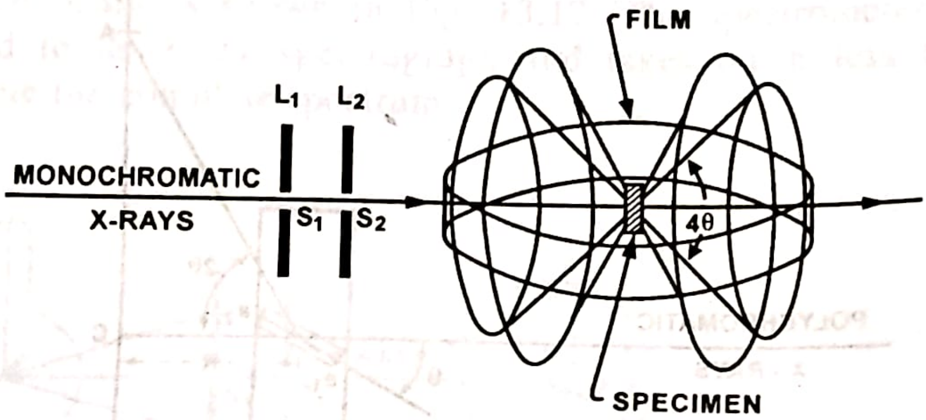


Fig. 13.19

The underlying principle of the powder technique is that since millions of tiny crystals in the powder have completely random orientations, all possible diffraction planes will be available for Bragg reflection to take place. Such reflections will take place from many sets of parallel planes lying at different angles to the incident beam of X-rays. Moreover, each set will give not only the first order reflections, but reflections of the higher orders as well. Since all orientations are equally likely, the reflected rays form a cone when a beam of X-rays pass through this powder. The axis of the cone lies along the direction of the incident beam and the semi-vertical angle of the cone is twice the glancing angle for that particular set of planes. For each set of planes and for each order, there will be such a cone of reflected X-rays. The intersection of these rings with a photographic film, set with its plane normal to the incident beam, form a series of concentric circular rings. Radii of these rings are recorded and the film can be used to find the glancing angle and hence the interplanar spacing of the crystalline substance.

Experimental arrangement

The experimental arrangement for powder crystal method is shown in Fig. 13.20. X-rays from a tube are allowed to pass through a filter F which absorbs all but one wavelength of the incident X-rays. This makes the X-ray monochromatic which is then collimated by two fine slits S_1 and S_2 cut in two lead plates L_1 and L_2. The collimated

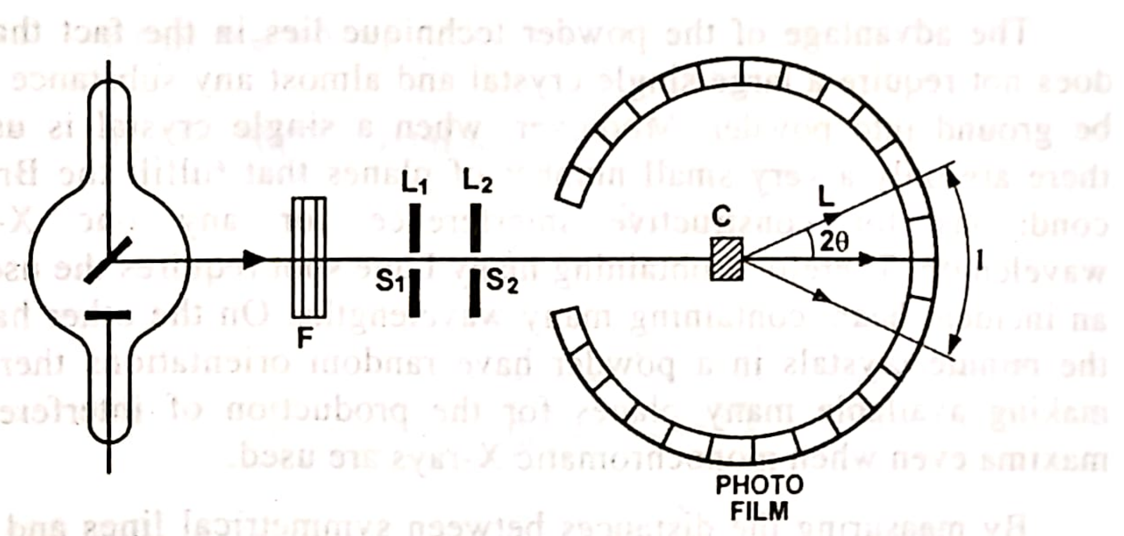


Fig. 13.20

beam falls on the powder specimen C suspended along the axis of the cylindrical Debye – Scherrer camera. The film fits round the inner surface of the camera and almost completely surrounds C in order to receive rays diffracted upto 180°. The film is of narrow width and hence only parts of the circular rings are registered on the film, their appearance being somewhat like that shown in Fig. 13.21 when the film is laid flat. The curvature of the arcs reverses after the angle of diffraction exceeds 90°.

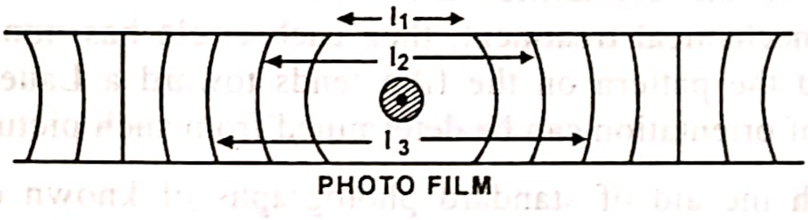


Fig. 13.21

Let the distances between symmetrical lines on the stretched photograph be l_1, l_2, l_3, etc. and the diameter of the cylindrical film is D

Then

$$\frac{l_1}{\pi D} = \frac{4\theta_1}{360^\circ}$$

$$\text{or, } \theta_1 = \frac{90^\circ}{\pi D}.l_1 = Kl_1 \qquad \text{where } K = \frac{90^\circ}{\pi D}$$

Similarly, $\theta_2 = Kl_2$ and $\theta_3 = Kl_3$ etc. Using these values of θ in Bragg's formula, interplanar spacing d can be calculated.

The advantage of the powder technique lies in the fact that it does not require a large single crystal and almost any substance can be ground into powder. Moreover, when a single crystal is used, there are only a very small number of planes that fulfill the Bragg conditions for constructive interference for any one X-ray wavelength. Therefore, obtaining many Laue spot requires the use of an incident beam containing many wavelengths. On the other hand, the minute crystals in a powder have random orientations thereby making available many planes for the production of interference maxima even when monochromatic X-rays are used.

By measuring the distances between symmetrical lines and the intensities of the lines with the help of a diffractometer, the following informations can be obtained.

(i) Crystalline substance can be distinguished from an amorphous substance because while the former produces lines on a photographic film, the latter does not.

(ii) Orientation of crystals in an extruded or drawn material such as a wire can be studied. If the crystallites in the wire are actually random in their orientation, a true powder pattern of circle results. But if the crystallites are somewhat oriented, say after thermal or mechanical treatment, then each circle has non-uniform intensity and the pattern on the film tends toward a Laue pattern. The degree of orientation can be determined from such pictures.

(iii) With the aid of standard photographs of known chemical compounds, their presence may be detected in an unknown mixture.

(iv) In simple cases, the size of the unit cell and the type of lattice can be determined.

Example 13.1 A beam of X-rays is incident on a NaCl crystal (lattice constant = 2.82 Å). The first order Bragg's reflection from (100) plane is observed at a grazing angle of 8° 35'. What is the wavelength of X-rays? At what angles would the second and third order Bragg reflection occur?

Soln.

According to Bragg's relation,

$2d \sin\theta = n\lambda.$

$$d_{100} \text{ for NaCl crystal} = \frac{1}{\left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right]^{1/2}}$$

$$= \frac{a}{\left(h^2+k^2+l^2\right)^{1/2}} \qquad (\because \text{ for NaCl } a=b=c)$$

$$= \frac{2.82\ \overset{\circ}{A}}{\left(1^2+0^2+0^2\right)^{1/2}} = 2.82\ \overset{\circ}{A}.$$

For first order Bragg reflection, $n = 1$ and $\theta = 8° 35'$

$\therefore\ 2d \sin\theta = n\lambda$

or, $2d \sin 8°35' = \lambda$

or, $\lambda = 2 \times 2.82\ \overset{\circ}{A} \times 0.1493 = 0.842\ \overset{\circ}{A}.$

For the second-order reflection, $n = 2$

$\therefore\ 2d \sin\theta = n\lambda$

or, $2 \times 2.82\ \overset{\circ}{A} \sin\theta = 2 \times 0.842\ \overset{\circ}{A}$

or, $\sin\theta = \dfrac{2\times 0.842\ \overset{\circ}{A}}{2\times 2.82\ \overset{\circ}{A}}$

or, $\theta = \sin^{-1}(0.2986) = 17°22'$

Similarly for third order reflection

$2d \sin\theta = n\lambda$

or, $\sin\theta = \dfrac{3\times 0.842\ \overset{\circ}{A}}{2\times 2.82\ \overset{\circ}{A}}$

$= 0.4478.$

or, $\theta = \sin^{-1}(0.4478)$

$= 26° 36'$

Example 13.2. *X-rays of wavelength 0.71 Å are reflected from the (100) face of a rock salt crystal whose lattice constant is 2.84 Å. Calculate the glancing angle corresponding to second order diffraction.*

Soln.

The inter-planar distance between (100) lattice planes is given by

$$D_{100} = \frac{1}{\left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right]^{1/2}} = \frac{a}{\left(h^2+k^2+l^2\right)^{1/2}} \quad (\because a=b=c)$$

$$= a = 2.84\ \text{Å}$$

$(\because h = 1, k = 0, l = 0)$

Now from Bragg's relation, we have

$$2d \sin\theta = n\lambda \quad \text{or,} \quad \sin\theta = \frac{n\lambda}{2d}$$

Here $d = 2.84\ \text{Å}$, $n = 2$, $\lambda = 0.71\ \text{Å}$

$$\therefore \sin\theta = \frac{2\times0.71}{2\times2.84} = 0.25$$

$$\therefore \theta = \sin^{-1}(0.25) = 14°\ 37'$$

Example 13.3. *Using a crystal of KCl, Bragg found the following glancing angles for first maxima from lattice planes defined by Miller indices of (100), (110) and (111) respectively*

$$\theta_1 = 5.22°,\ \theta_2 = 7.30° \text{ and } \theta_3 = 9.05°$$

Show that KCl has a simple cubic structure.

Soln.

From Bragg's relation, we have

$$2d \sin\theta = n\lambda$$

Here $n = 1$

$$\therefore \quad d = \frac{\lambda}{2\sin\theta}$$

For (100) plane $d_{100} = \dfrac{\lambda}{2\sin\theta_1}$

(110) plane $d_{110} = \dfrac{\lambda}{2\sin\theta_2}$

(111) plane $d_{111} = \dfrac{\lambda}{2\sin\theta_3}$

Then

$$d_{100} : d_{110} : d_{111} = \frac{\lambda}{2\sin\theta_1} : \frac{\lambda}{2\sin\theta_2} : \frac{\lambda}{2\sin\theta_3}$$

$$= \frac{1}{\sin\theta_1} : \frac{1}{\sin\theta_2} : \frac{1}{\sin\theta_3}$$

$$= \frac{1}{\sin 5.22^\circ} : \frac{1}{\sin 7.30^\circ} : \frac{1}{\sin 9.05^\circ}$$

$$= \frac{1}{0.091} : \frac{1}{0.127} : \frac{1}{0.157}$$

$$= 10.98 : 7.874 : 6.369$$

$$= 1 : 0.71 : 0.58.$$

This represents the characteristic of a simple cubic structure. Hence KCl has a simple cubic structure.

Example 13.4. *Calculate the longest wavelength that can be analysed by a rock salt crystal of spacing d = 2.82 Å*

(i) in the first order

(ii) in the second order

Soln.

According to Bragg's equation

$$2d \sin \theta = n\lambda$$

$$\text{or, } \lambda = \frac{2d \sin \theta}{n}$$

(i) $\theta = 90°$ for the longest wavelength

and for first order $n = 1$, $d = 2.82\ \text{Å}$

$$\therefore\ \lambda = \frac{2 \times 2.82 \times \sin 90°}{1} = 5.64\ \text{Å}$$

(ii) In the second order $n = 2$

$$\therefore\ \lambda = \frac{2 \times 2.82 \times \sin 90°}{2} = 2.82\ \text{Å}$$

Example 13.5. *For a certain b.c.c. crystal, the 110 planes have separation of 1.181 Å. These 110 planes are irradiated with X-rays of wavelength 1.540 Å. How many orders of Bragg reflections can be observed in this case?*

Soln.

According to Bragg's relation

$$2d \sin \theta = n\lambda$$

where n is the order of reflection, λ = the wavelength of incident X-rays and d = inter-planer separation.

Now the maximum possible value of $\sin\theta = 1$

$$\therefore\ 2d = n\lambda$$

$$\text{or, } n = \frac{2d}{\lambda}$$

$$= \frac{2 \times 1.181}{1.540}$$

$$= 1.54.$$

Since fraction is meaningless, only first order reflection is possible.

Example 13.6. Calculate the glancing angle on the cube (110) of a rock salt crystal ($a = 2.81\,\mathring{A}$) corresponding to second order diffraction maximum for the X-rays of wavelength $0.71\,\mathring{A}$.

Soln.

From the Bragg's relation

$2d \sin\theta = n\lambda$ we have

$$\sin\theta = \frac{n\lambda}{2d}$$

here $n = 2$

$\lambda = 0.71\,\mathring{A}$

and $d = \frac{a}{\sqrt{h^2 + k^2 + l^2}} = \frac{2.81}{\sqrt{1^2 + 1^2}} = \frac{2.81}{\sqrt{2}}\,\mathring{A} = 1.987\,\mathring{A}$

$$= \frac{2 \times 0.71\,\mathring{A}}{2 \times 1.9871\,\mathring{A}}$$

$$= 0.3573$$

$$\therefore\ \theta = \sin^{-1}(0.3573)$$

$$= 20^\circ\, 54'.$$

Example 13.7. X-rays of wavelength $0.36\,\mathring{A}$ are diffracted in first order at an angle of 4.80 in Bragg's crystal spectrometer. Find the effective spacing of atomic layers in the crystal.

Soln.

From the Bragg's relation, we have

$$2d \sin\theta = n\lambda$$

$$\text{or, } d = \frac{n\lambda}{2\sin\theta}$$

Here $\theta = 4.80^\circ$, or, $\sin\theta = 0.084$

$\lambda = 0.36\,\mathring{A}$ $\qquad n = 1$

$$\therefore\ d = \frac{n\lambda}{2\sin\theta} = \frac{1 \times 0.36\,\mathring{A}}{2 \times 0.084} = 2.15\,\mathring{A}.$$

Example 13.8. Calculate glancing angle at which the first and second order diffraction maxima will be observed when X-rays of 1.92 Å wavelength are reflected from a cleavage plane of calcite for which d = 3.029 Å.

Soln.

Bragg's relation for X-ray diffraction is given by

$$2d \sin \theta = n\lambda$$

where d = inter-planer separation of lattice planes

n = order of diffraction maximum

λ = wavelength of incident X-rays

and θ = glancing angle at which diffraction maximum is observed.

In the first order, we have n = 1

$$\therefore \quad \sin \theta = \frac{n\lambda}{2d} = \frac{1 \times 1.92\,\text{Å}}{2 \times 3.029\,\text{Å}} = 0.3169$$

$$\therefore \quad \theta = \sin^{-1}(0.3169) = 18°29'$$

For the second order, n = 2

$$\therefore \quad \sin \theta = \frac{n\lambda}{2d} = \frac{1 \times 1.92\,\text{Å}}{2 \times 3.029\,\text{Å}} = 0.6338$$

$$\therefore \quad \theta = \sin^{-1}(0.6338) = 39°20'.$$

Example 13.9. The radiation from an X-ray tube operated at 50KV are diffracted by a cubic KCl crystal of molecular mass 74.6 and density 1.99 x 10^3 kg/m^3. Calculate (i) the short wavelength limit of the spectrum from the tube and (ii) glancing angle for first-order reflection from the principal planes of the crystal for that wavelength.

Soln.

(i) $\lambda_{min} = \dfrac{12400}{50\times10^3} = 0.248\,\mathring{A}$.

(ii) From Bragg's relation, we have

$$2d \sin \theta = n\lambda$$

Here $\lambda = 0.248\,\mathring{A}$, $n = 1$

Now d can be obtained from the following relation

$$a^3 \rho = \frac{nM}{N}$$

where $\rho = 1.99 \times 10^3$ kgm^3

KCl belongs to fcc crystal system. $\therefore$ n, the number of atoms per unit cell = 4.

M = 74.6 and N, the Avogadro's number = 6.02×10^{26} per kg-mole.

$$\therefore \quad a^3 \times 1.99 \times 10^3 = \frac{4\times74.6}{6.02\times10^{26}} \quad \text{or,} \quad a^3 = 249 \times 10^{-30}$$

$$\text{or,} \quad a = 6.29 \times 10^{-10}$$

For a fcc crystal the separation between the (100) planes,

$$\therefore \quad d = \frac{a}{2} = \frac{6.28\times10^{-10}}{2} = 3.14 \times 10^{-10}\text{m} = 3.14\,\mathring{A}.$$

Therefore, from

$2d \sin \theta = n\lambda$ we have

$$\sin \theta = \frac{n\lambda}{2d} = \frac{1\times0.248\,\mathring{A}}{2\times3.14\,\mathring{A}} = 0.0395$$

$$\therefore \quad \theta = \sin^{-1}(0.0395) = 2^\circ\, 12'$$

Example 13.10. *In a Bragg's spectrometer, the glancing angle for first order spectrum was observed to be 8°. Calculate the wavelength of X-rays if d = 2.82 x 10^{-10}m. At what angle will the second maximum occur?*

Soln.

$$2d \sin \theta = n\lambda$$

$$\lambda = \frac{2d \sin \theta}{n}$$

(i) Here n = 1, θ = 8° ∴ sin θ = 0.1392

$d = 2.82 \times 10^{-10}$m

$$\therefore \lambda = \frac{2 \times 2.82 \times 10^{-10} \times 0.1392}{1}$$

$$= 0.785 \times 10^{-10} \text{m.}$$

(ii) For the second order n = 2,

$$\therefore \sin \theta = \frac{n\lambda}{2d} = \frac{2 \times 0.785 \times 10^{-10}}{2 \times 2.82 \times 10^{-10}} = 0.2783$$

$$\therefore \theta = \sin^{-1}(0.2783) = 16.15°.$$

Example 13.11. *A set of lattice planes reflects X-rays of wavelength 1.32 Å at a glancing angle of 9°30'. Deduce the possible spacing for this set of planes.*

Soln.

we have

$$2d \sin \theta = n\lambda$$

$$\text{or, } d = \frac{n\lambda}{2 \sin \theta}$$

Here θ = 9° 30' ∴ sin θ = 0.1650

$$\lambda = 1.32 \text{ Å} = 1.32 \times 10^{-10} \text{m}$$

$$\therefore \quad d = \frac{n\times1.32\times10^{-10}}{2\times0.1650}$$

$$= n \times 4 \times 10^{-10}\text{m.}$$

where n = 1, 2, 3, 4, ...

Hence the possible spacings are

4×10^{-10}m, 8×10^{-10}m, 12×10^{-10}m and so on *i.e.*, 4 Å, 8 Å, 12 Å or integral multiples of 4 Å.

Example 13.12. Compute the wavelength of the most energetic photons emitted by an X-ray tube operated at a steady potential of 80,000 volts. At what glancing angle would these photons be reflected, in the first order from the 100 planes of sodium?

Soln.

If ν_{max} is the frequency of the most energetic photons emitted by the X-ray tube then,

$$eV = h\nu_{max} = \frac{hc}{\lambda_{min}}$$

where λ_{min} is the corresponding wavelength.

Now $h = 6.625 \times 10^{-34}$ J-s

$c = 3 \times 10^{8}$ ms^{-1}

$e = 1.6 \times 10^{-19}$ coul

and V = 80,000 V.

$$\therefore \quad \lambda_{min} = \frac{hc}{eV} = \frac{6.625\times10^{-34}\times3\times10^{8}}{1.6\times10^{-19}\times80{,}000}$$

$$= 0.1553 \times 10^{-10}\text{m}$$

$$= 0.1553\ \text{Å}.$$

According to Bragg's law,

$$2d \sin \theta = n\lambda.$$

where θ is the glancing angle at which the photons are reflected, n being the order of reflection

For 100 planes of sodium, $d = 2.81 \times 10^{-10}$ and $\lambda = 0.1553 \times 10^{-10}$m.

$\therefore \quad 2 \times 2.81 \times 10^{-10} \sin\theta = 0.1553 \times 10^{-10} \times 1$

$$\text{or, } \sin\theta = \frac{0.1553\times10^{-10}}{2\times2.81\times10^{-10}} = 0.02763$$

$$\text{or, } \theta = \sin^{-1}(0.02763)$$

$$= 1^\circ\, 35'.$$

Example 13.13. *A powder pattern is obtained for lead with radiations of $\lambda = 1.54$ Å. The (220) reflection is observed at Bragg angle $\theta = 32°$. What is the lattice parameter of lead and the radius of the atom?*

Soln.

Let us assume that the reflection in question is of the first order, *i.e.*, n = 1.

So from $2d \sin\theta = n\lambda$ we have

$$d = \frac{n\lambda}{2\sin\theta} = \frac{1\times1.54\times10^{-10}}{2\sin 32^\circ}$$

$$= \frac{1.54\times10^{-10}}{2\times0.53} = 1.45 \times 10^{-10}\text{m}.$$

The interplanar spacing and the lattice constant of lead are connected by the relation

$$d = \frac{a}{\sqrt{h^2+k^2+l^2}}$$

$$\text{or, } a = d.\sqrt{h^2+k^2+l^2} = 1.45 \times 10^{-10}\sqrt{2^2+2^2+0^2}$$

$$= 4.1 \times 10^{-10}\text{m}.$$

Now, lead has face-centred cubic (FCC) lattice and for FCC lattice crystals

$$a = 4r / \sqrt{4}$$

$$\text{or,} \quad r = \frac{a.\sqrt{2}}{4} = \frac{\sqrt{2}\times 4.1\times 10^{-10}}{4}$$

$$= 1.45 \times 10^{-10}\text{m.}$$

***Example 13.14.** If the potential difference applied across an X-ray tube is 5KV and the current through it is 2mA, calculate*

(i) the number of electrons striking the target per second and

(ii) the speed at which they strike it.

$e = 1.6 \times 10^{-19}$ C and $m = 9.1 \times 10^{-31}$ kg.

Soln.

(i) If n is the number of electrons striking the anode per second, then

$$n.e = I$$

$$\text{or,} \quad n = \frac{I}{e} = \frac{2\times 10^{-3}}{1.6\times 10^{-19}} = 1.25 \times 10^{16} \text{ per second.}$$

(ii) As shown in Art. 13.3

$$\frac{1}{2}mv^2 = eV, \quad \text{or,} \quad v = \sqrt{\frac{2eV}{m}}$$

Substituting the values of n and e, we get

$$v = 5.93 \times 10^5 \sqrt{V} \quad \text{ms}^{-1}$$

Hence the velocity of the electrons for a potential difference of 5 KV,

$$v = 5.93 \times 10^5 \sqrt{5000} = 4.15 \times 10^7 \text{ m s}^{-1}.$$

***Example 13.15.** An X-ray tube passes 5mA at a potential difference of 100 KV. Calculate the maximum speed of the electrons*

striking the target and the rate of production of heat at the target if only 0.1 per cent of the incident energy is converted into X-radiations. J = 4.18 joules per calorie.

Soln.

(i) $v = 5.93 \times 10^5 \sqrt{V}$ m/s (Art. 12.1)

$= 5.93 \times 10^5 \sqrt{100,000}$ m/s

$= 1.88 \times 10^8$ m/s.

(ii) Incident power = voltage x current

$= 100,000 \times 5 \times 10^{-3}$

$= 500$ W.

Power converted into heat = 99.9% of 500 W

= 499.5 W

= 499.5 joules per sec.

∴ Heat produced per second

$$= \frac{499.5}{4.18} = 119 \text{ cal/s.}$$

***Example 13.16.** An X-ray tube operated at 30 KV emits a continuous X-ray spectrum with a short wavelength limit $\lambda_{min} = 0.414$ Å. Calculate Planck's constant h if $e = 1.602 \times 10^{-19}$ C and $c = 3 \times 10^8$ m/s.*

Soln.

$$\lambda_{min} = \frac{ch}{eV} \quad \text{(Art. 13.3)}$$

$$h = \frac{eV\lambda_{min}}{c}$$

$$= \frac{1.602\times10^{-19}\times30\times10^3\times0.414\times10^{-19}}{3\times10^8}$$

$= 6.63 \times 10^{-34}$ J-s.

***Example 13.17.** An impure platinum target emits a strong L_α line of wavelength 1.321 Å and a weak L_α line of wavelength 4.174 Å. Identify the impurity taking the nuclear screening constant to be 7.4 for L_α lines. Atomic number for tungsten is 78.*

Soln.

According to Moseley's law

$$\sqrt{\nu} = a\,(z = \sigma)$$

Let ν_1 and ν_2 be the frequencies of the L_α lines of platinum and the impurity respectively. Let their atomic number be Z_1 and Z_2 respectively. Then

$$\sqrt{\nu_1} = a\,(Z_1 - \sigma) \quad \text{and} \quad \sqrt{\nu_2} = a\,(Z_2 - \sigma)$$

$$\text{or,} \quad \sqrt{\frac{\nu_1}{\nu_2}} = \frac{a\,(z_1 - \sigma)}{a\,(z_2 - \sigma)} = \frac{z_1 - \sigma}{z_2 - \sigma}$$

$$\text{Now } n = \frac{c}{\lambda}; \text{ hence } \frac{\nu_1}{\nu_2} = \frac{\lambda_2}{\lambda_1}$$

$$\therefore \quad \sqrt{\frac{\nu_1}{\nu_2}} = \sqrt{\frac{\lambda_1}{\lambda_2}} = \frac{z_1 - \sigma}{z_2 - \sigma}$$

$$\text{or,} \quad Z_2 - \sigma = (Z_1 - \sigma)\sqrt{\frac{\lambda_1}{\lambda_2}}$$

Now $\sigma = 7.4$; $\lambda_1 = 1.321$ Å and $\lambda_2 = 4.174$ Å and $Z_1 = 78$

$$\therefore \quad Z_2 = 7.4 + (78 - 7.4)\sqrt{\frac{1.321}{4.174}}$$

$$= 47.12.$$

The atomic number of the impurity is 47.12 which, of course, is silver.

13.13 Compton Effect

In 1923 A.H. Compton, while making a spectroscopic study of X-rays scattered by matter, discovered that, when a beam of X-rays of well defined wavelength λ is scattered through an angle θ by sending the radiation through the metallic foil, the scattered radiation, in addition to the expected incident wavelength λ contains a component of well-defined wavelength λ' which is longer than λ. This phenomenon is called the *Compton effect*.

The principle of the experiment performed by Compton is shown in Fig. 13.22 X-rays from a molybdenum target is allowed to pass through a filter which absorbs all wavelengths except MoK_α line. The MoK_α line of wavelength $\lambda = 0.71\ \mathring{A}$ is then allowed to fall on a carbon target. The radiation scattered from the target at various

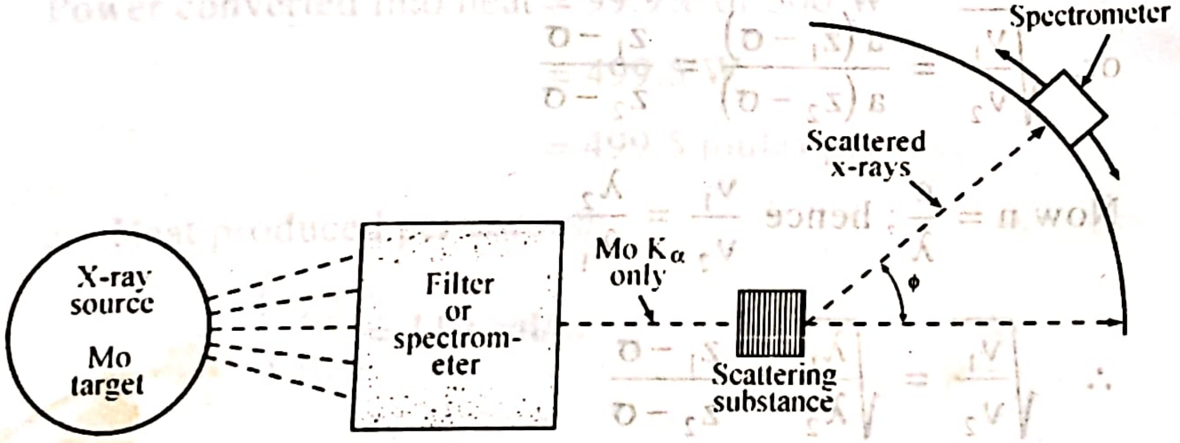


Fig. 13.22

scattering angles φ was analyzed by a spectrometer in which the grating used was a crystal (Bragg spectrometer). The results are shown in Fig. 13.23.

At $\varphi = 0$, the spectrum of the scattered radiation is essentially the spectrum of the radiation incident on the carbon target. At other scattering angles, two peaks can be seen in the spectrum of the scattered radiation. One is called the *unmodified* line, having the same wavelength as the incident radiation; the other, which is broader, is called the *modified* line. The wavelength of the modified line is longer that that of the incident radiation and the wavelength increment increases with the increase of the scattering angle ϕ. When the angle between the incident and the scattered radiation was

90°, the difference in wavelength was found to be 0.0236×10^{-8} cm, independent of the wavelength of the primary (incident) beam and of the nature of the scattering material.

The unmodified peak shown in Fig. 13.23 can be explained in terms of classical electrodynamics without resorting to the concept of photons as carriers of the electromagnetic radiation. To account for the presence of the shifted component and for the amount of the

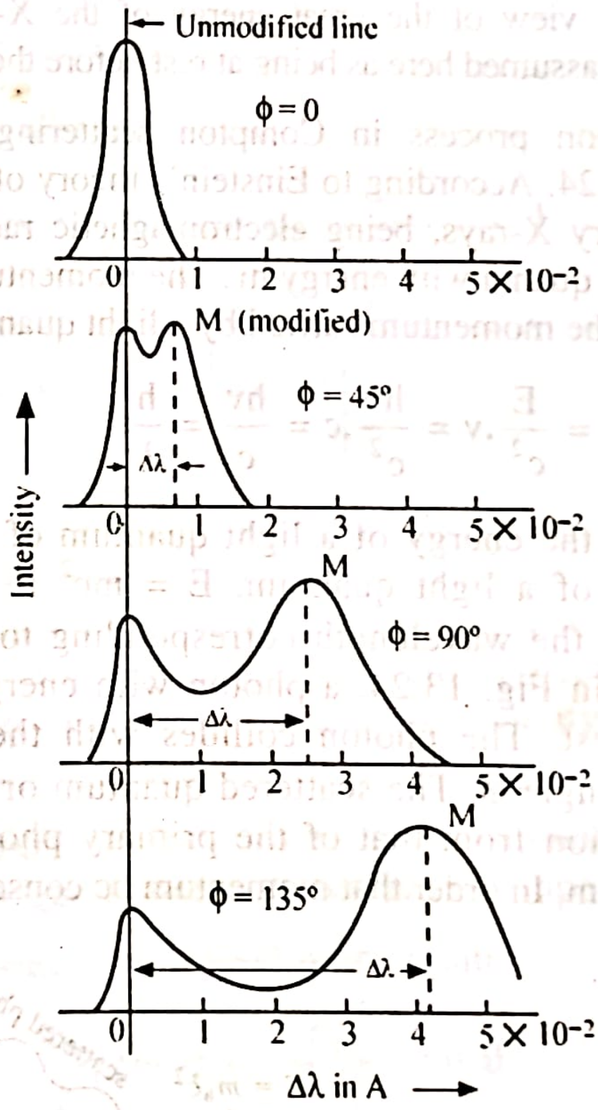


Fig. 13.23

shift, Compton showed that the *corpuscular* theory of X-rays could explain this peak. The scattering of X-rays without any change in wavelength (unmodified line) is referred to as *coherent* scattering while the scattering with change in wavelength (modified line) is called *incoherent* scattering. The incoherent scattering is often called the *Compton scattering*.

In order to account for change in wavelength of the incident X-rays, Compton assumed that the scattering process could be treated as an elastic collision between a photon and a free electron in which both energy and momentum are conserved. The electron which is alone effective in this act of scattering is treated as free although it is bound to the atomic configuration. This is justified if the binding energy of the electron is negligible compared to the energy of the incident photon. Thus, in Compton scattering, the electrons – especially the more loosely bound electrons can be treated as free in view of the great energy of the X-ray spectrum. The electron is further assumed here as being at rest before the collision.

The collision process in Compton scattering is schematically shown in Fig. 13.24. According to Einstein's theory of the photo-electric effect, the primary X-rays, being electromagnetic radiations like light, are propagated as quanta with energy $h\nu$. The momentum associated with this energy (*i.e.*, the momentum carried by a light quantum or photon) is

$$p = mv = \frac{E}{c^2}.v = \frac{h\nu}{c^2}.c = \frac{h\nu}{c} = \frac{h}{\lambda}$$

where $E = h\nu$ is the energy of a light quantum of frequency ν, c is the velocity (v) of a light quantum, $E = mc^2$ is the mass energy relation and λ is the wavelength corresponding to the frequency ν. Then, as shown in Fig. 13.24, a photon with energy $h\nu$ approaches an electron at rest. The photon collides with the electron and is deflected by an angle φ. The scattered quantum or photon moves in a different direction from that of the primary photon and carries a different momentum. In order that momentum be conserved, the electron

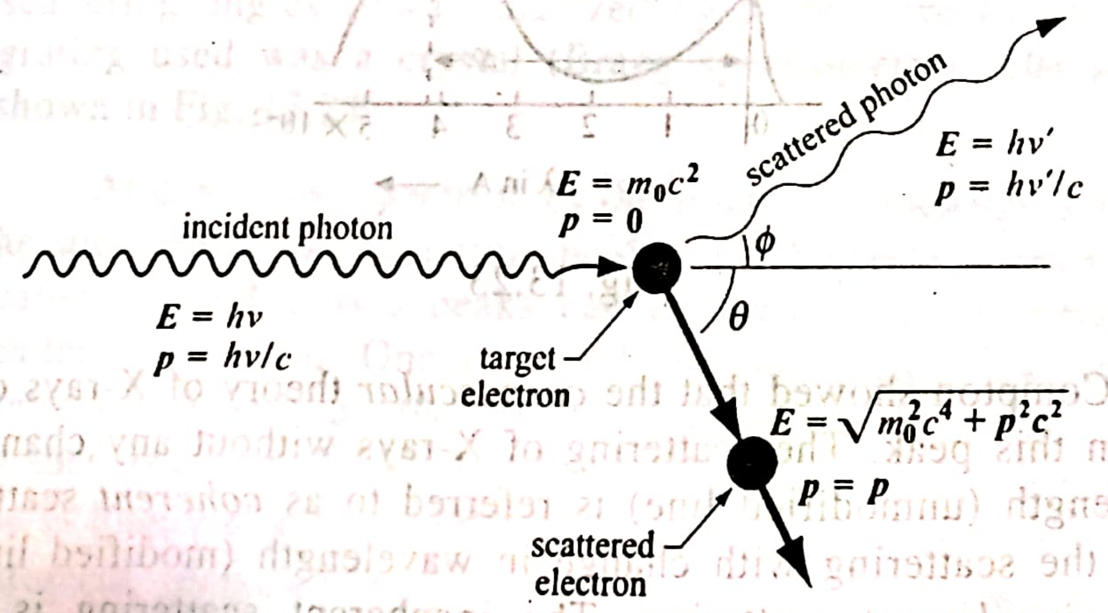


Fig. 13.24

which scatters the photon must recoil with a momentum equal to the vector difference between the momentum of incident photon and the momentum of the scattered photon. The energy of this *recoiling* electron must come from the primary photon. Thus the scattered photon has less energy – and hence a lower frequency or longer wavelength – than that of the primary photon. Thus

loss in photon energy = gain in electron energy

$$\text{or, } h\nu - h\nu' = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}}.c^2 - m_oc^2$$

or, the requirement that energy be conserved gives

$$h\nu = h\nu' + m_oc^2\left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right] \qquad \text{(i)}$$

where ν is the frequency of the incident photon, ν' that of the scattered photon, m_oc^2 is the rest mass energy of the electron and v is the recoil velocity of the electron. In eqn. (i) relativistic kinetic energy of the electron is used because the velocity of the electron may be great enough for relativistic effects to be significant.

The requirement for the conservation of momentum gives two equations. In the original photon direction (x-component),

initial momentum = final momentum

$$\frac{h\nu}{c} + 0 = \frac{h\nu'}{c}\cos\varphi + \frac{m_o v}{\sqrt{1-\frac{v^2}{c^2}}}\cos\theta \qquad \text{(ii)}$$

and perpendicular to this direction (y-component)

initial momentum = final momentum

$$0 = \frac{h\nu'}{c}\sin\varphi - \frac{m_o v}{\sqrt{1-\frac{v^2}{c^2}}}\sin\theta \qquad \text{(iii)}$$

φ is the angle between the directions of the initial and scattered photons, and θ is that between the directions of the initial photon and recoil electron. To solve these equations, let us write $\beta = \frac{v}{c}$. Eqns. i, ii and iii then become

$$h\nu = h\nu' + m_0c^2\left[\frac{1}{\sqrt{1-\beta^2}} - 1\right] \tag{iv}$$

$$\frac{h\nu}{c} = \frac{h\nu'}{c}\cos\varphi + \frac{m_0\beta c}{\sqrt{1-\beta^2}}\cos\theta \tag{v}$$

$$0 = \frac{h\nu'}{c}\sin\varphi - \frac{m_0\beta c}{\sqrt{1-\beta^2}}\sin\theta \tag{vi}$$

For a particular angle of scattering ϕ, equations iv, v and vi contain three unknowns, namely v, β and θ, the expressions for which can be found by a straight forward solution of these equations. For comparison with experiment, we are, however, more interested in the shift in the wavelength. Therefore, in terms of wavelength, we write

$$\lambda = \frac{c}{\nu} \quad \text{and} \quad \lambda' = \frac{c}{\nu'}$$

for the incident and scattered radiations respectively. Eqns. iv, v and vi then become

$$\frac{h}{\lambda} - \frac{h}{\lambda'} + m_0c = \frac{m_0c}{\sqrt{1-\beta^2}} \tag{vii}$$

$$\frac{h}{\lambda} - \frac{h}{\lambda'}\cos\varphi = \frac{m_0\beta c}{\sqrt{1-\beta^2}}\cos\theta \tag{viii}$$

$$\frac{h}{\lambda'}\sin\varphi = \frac{m_0\beta c}{\sqrt{1-\beta^2}}\sin\theta \tag{ix}$$

Squaring equations viii and ix and adding, we have

$$\frac{h^2}{\lambda^2} + \frac{h^2}{\lambda'^2} - \frac{2h^2\cos\varphi}{\lambda\lambda'} = \frac{m_o^2\beta^2c^2}{1-\beta^2}$$

$$= \frac{m_o^2c^2}{1-\beta^2} - m_o^2c^2 \qquad \text{(x)}$$

Squaring eqn. (vii), we get

$$\frac{h^2}{\lambda^2} + \frac{h^2}{\lambda'^2} - \frac{2h^2}{\lambda\lambda'} + 2m_och\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) + m_o^2c^2 = \frac{m_o^2c^2}{1-\beta^2} \qquad \text{(xi)}$$

Subtracting (x) from (xi), we get

$$\frac{2h^2}{\lambda\lambda'}(\cos\varphi - 1) + 2m_och\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = 0$$

$$\text{or,}\quad 2m_och\left(\frac{\lambda' - \lambda}{\lambda\lambda'}\right) = -\frac{2h^2}{\lambda\lambda'}(\cos\varphi - 1)$$

$$= \frac{2h^2}{\lambda\lambda'}(1 - \cos\varphi)$$

$$\text{or,}\quad \Delta\lambda = \lambda' - \lambda = \frac{h}{m_0c}(1 - \cos\varphi) \qquad (13.13)$$

Eqn. 13.13 was derived by Arthur H. Compton in the early 1920s. The phenomenon described by this equation was first observed by him and is known as Compton effect. The equation constitutes very strong evidence in support of the quantum theory of radiation.

Eqn. 13.13 gives the change in the wavelength of a photon that is scattered through φ by a particle of rest mass m_o; it is independent of the wavelength of the incident photon but varies with the scattering angle ϕ as $(1 - \cos\varphi)$.

The quantity h/m_oc has the dimension of length. If the values $h = 6.624 \times 10^{-27}$ erg-s, $m = 9.107 \times 10^{-28}$ gm and $c = 2.998 \times 10^{10}$ cm/sec are inserted, then

$$\Delta\lambda = 0.0242 \times 10^{-8}(1 - \cos\varphi)\ \text{cm}$$

$$= 0.0242(1 - \cos\varphi)\ \text{Å} \qquad (13.14)$$

Thus eqn. 13.14 tells us that for a given angle of scattering φ, the wavelength of the scattered radiation should be greater than that of the incident radiation by the amount $0.0242\,(1-\cos\varphi)\,\mathring{A}$.

(i) If $\varphi = 90°$, $\cos\varphi = 0$ and hence $\Delta\lambda = 0.0242\,\mathring{A}$ which agrees very well with the observed value of $0.0236\,\mathring{A}$. The quantity $h/m_o c$ ($= 0.0242\,\mathring{A}$) is called the *Compton wavelength.*

(ii) If $\varphi = 0$, $\cos\varphi = 1$ and hence $\Delta\lambda = 0$.

(iii) When $\varphi = 180°$, $\cos\varphi = -1$ and hence $\Delta\lambda = \dfrac{2h}{m_o c} = 0.0484\,\mathring{A}$.

Hence the greatest change in wavelength that can occur takes place when $\varphi = 180°$ when the wavelength change will be twice the Compton wavelength.

The Compton wavelength will be considerably less for other particles owing to their larger rest masses. Hence the maximum wavelength change in Compton effect is $0.0484\,\mathring{A}$ which takes place for an electron when the scattering angle is 180°. Changes of this magnitude are readily observable only in X-rays since the shift in wavelength for visible light is less than 0.01 per cent of the initial wavelength while for X-rays of $\lambda = 1\,\mathring{A}$ it is several per cent. The predicted dependence of $\Delta\lambda$ on the scattering angle φ was also verified by experiment.

In the experimental investigation, the scattered X-rays at each angle of scattering was found to include a substantial proportion having the initial wavelength. This is not hard to understand. In deriving equation (13.13), the scattering particle was assumed to be free to move at will, a reasonable assumption since many of the electrons in matter are only loosely bound to their parent atoms. Other electrons, however, are very tightly bound and when struck by a photon, the entire atom recoils instead of the single electron. In this event the value of m_o, as used in eqn. 13.13, is that of the entire atom, usually tens of thousands of times greater than that of an electron. The resulting Compton shift is accordingly so minute as to be detectable. Further, the broadening of the

modified or Compton peak is the result of the motion of the electron in the atom. Eqn. 13.13 was derived on the basis of an electron initially at rest. If the motion of the electron is taken into account, the broadening of the modified peak can be completely explained.

Kinetic energy of the recoil electron

The kinetic energy of the recoil electron may be obtained from the relation

$$E_r = h\nu - h\nu'$$

where $h\nu$ is the energy of the incident photon and $h\nu'$ is that of the scattered photon.

From eqn. 13.13 the change in wavelength in a Compton scattering is given by

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{m_o c}(1 - \cos\varphi)$$

$$\text{or,} \quad \frac{c}{\nu'} - \frac{c}{\nu} = \frac{h}{m_o c}(1 - \cos\varphi)$$

$$\frac{1}{\nu'} - \frac{1}{\nu} = \frac{h}{m_o c^2}(1 - \cos\varphi)$$

$$\frac{1}{h\nu'} - \frac{1}{h\nu} = \frac{1}{m_o c^2}(1 - \cos\varphi)$$

$$\frac{1}{h\nu'} = \frac{1}{h\nu} + \frac{1}{m_o c^2}(1 - \cos\varphi)$$

$$= \frac{m_o c^2 + h\nu(1-\cos\varphi)}{h\nu . m_o c^2}$$

$$\text{or,} \quad h\nu' = \frac{h\nu . m_o c^2}{m_o c^2 + h\nu(1-\cos\varphi)}$$

$$= \frac{h\nu}{1 + \frac{h\nu}{m_o c^2}(1-\cos\varphi)}$$

$$= \frac{h\nu}{1+\alpha(1-\cos\varphi)}$$

$$\text{where } \alpha = \frac{h\nu}{m_0 c^2}$$

Hence the kinetic energy of the recoil electron is

$$E_r = h\nu - h\nu'$$

$$= h\nu - \frac{h\nu}{1+\alpha(1-\cos\varphi)}$$

$$= h\nu\left[1-\frac{1}{1+\alpha(1-\cos\varphi)}\right]$$

$$= h\nu\left[\frac{1+\alpha(1-\cos\varphi)-1}{1+\alpha(1-\cos\varphi)}\right]$$

$$= h\nu\left[\frac{\alpha(1-\cos\varphi)}{1+\alpha(1-\cos\varphi)}\right]$$

$$= h\nu\left[\frac{\alpha.2\sin^2\frac{\varphi}{2}}{1+\alpha.2\sin^2\frac{\varphi}{2}}\right]$$

$$= h\nu\left[\frac{2\alpha\sin^2\frac{\varphi}{2}}{1+2\alpha\sin^2\frac{\varphi}{2}}\right] \qquad (13.15)$$

Direction of recoil electron

From eqns. ii and iii, we have

$$mvc\cos\theta = h\nu - h\nu'\cos\varphi \qquad \text{(i)}$$

$$mvc\sin\theta = h\nu'\sin\varphi \qquad \text{(ii)}$$

Dividing ii by i, we get

$$\tan\theta = \frac{h\nu'\sin\varphi}{h\nu - h\nu'\cos\varphi}$$

$$= \frac{\nu'\sin\varphi}{\nu - \nu'\cos\varphi}$$

$$\text{But } \nu' = \frac{\nu}{1+\alpha(1-\cos\varphi)} = \frac{\nu}{1+2\alpha\sin^2\frac{\varphi}{2}}$$

$$\therefore \tan\theta = \frac{\nu\sin\varphi/(1+2\alpha\sin^2\frac{\varphi}{2})}{\nu - \dfrac{\nu\cos\varphi}{1+2\alpha\sin^2\frac{\varphi}{2}}}$$

$$= \frac{\sin\varphi}{1+2\alpha\sin^2\frac{\varphi}{2} - \cos\varphi}$$

$$= \frac{\sin\varphi}{(1-\cos\varphi)+2\alpha\sin^2\frac{\varphi}{2}}$$

$$= \frac{2\sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\varphi}{2}\right)}{2\sin^2\frac{\varphi}{2}+2\alpha\sin^2\frac{\varphi}{2}}$$

$$= \frac{\cos\left(\frac{\varphi}{2}\right)}{\sin\frac{\varphi}{2}+\alpha\sin\frac{\varphi}{2}}$$

$$= \frac{\cos\frac{\varphi}{2}}{(1+\alpha)\sin\frac{\varphi}{2}}$$

$$= \frac{\cot\frac{\varphi}{2}}{1+\alpha}$$

$$\text{But } \alpha = \frac{h\nu}{m_o c^2}$$

$$\therefore \quad \tan\theta = \frac{\cot\frac{\varphi}{2}}{1 + h\nu / m_o c^2}$$

$$= \frac{\cot\frac{\varphi}{2}}{1 + \frac{h}{m_o c\lambda}} \qquad (\because \nu = \frac{c}{\lambda}) \tag{13.16}$$

Eqn. 13.16 shows that the recoil angle θ depends on the scattering angle φ. If $\varphi = 0°$, $\theta = 90°$ and when $\varphi = 180°$, $\theta = 0°$. Thus as φ varies from 0 to 180°, θ varies from 90° to 0°. This indicates that electron can get recoiled only in the forward direction at angles less than 90°, whereas a photon can get scattered in all directions.

Example 13.18. *Find the change in wavelength of an X-ray photon when it is scattered through an angle of 90° by a free electron.*

Soln.

$$\Delta\lambda = \frac{h}{m_o c}(1 - \cos\phi)$$

Here $\phi = 90°$ or, $\cos 90° = 0$

$\therefore$ $h = 6.624 \times 10^{-34}$ J-s

$m_o = 9.1 \times 10^{-31}$ Kg

$c = 3 \times 10^8$ m/sec.

$$\therefore \quad \Delta\lambda = \frac{h}{m_o c} = \frac{6.624\times10^{-34}\times10^{-19}}{9.1\times10^{-31}\times3\times10^8}$$

$$= 0.0242 \times 10^{-10}\text{m}$$

$$= 0.0242 \text{ Å}.$$

***Example 13.19.** Monochromatic X-rays of wavelength 0.124 Å are scattered by a carbon block. Find the wavelength of X-rays scattered through 180°.*

Soln.

Change in wavelength

$$\Delta\lambda = \frac{h}{m_o c}(1 - \cos\phi)$$

Here $\phi = 180°$ or, $\cos 180° = -1$

$$\therefore \Delta\lambda = \frac{2h}{m_o c} = \frac{2 \times 6.624 \times 10^{-34}}{9.1 \times 10^{-31} \times 3 \times 10^{8}} = 0.0484 \text{ Å}$$

Now $\Delta\lambda = \lambda' - \lambda$

or, $\lambda = \lambda + \Delta\lambda$

$= 0.124 + 0.0484$

$= 0.1724$ Å.

***Example 13.20.** An X-ray photon collides with an electron at rest and is scattered through 90°. What is its frequency after collision? Its initial frequency is 3×10^{19} Hz.*

Soln.

Wavelength of the photon, $\lambda = \frac{c}{\nu} = \frac{3 \times 10^{8}}{3 \times 10^{19}}$

$= 10^{-11} = 0.1 \times 10^{-10} = 0.1$ Å

Change in wavelength

$$\Delta\lambda = \frac{h}{m_o c}(1 - \cos\phi)$$

$$= \frac{h}{m_o c}(1 - \cos 90°)$$

$$= \frac{h}{m_o c} = \frac{6.624 \times 10^{-34}}{9.1 \times 10^{-31} \times 3 \times 10^{8}} = 0.0242 \text{ Å}$$

Now $\Delta\lambda = \lambda' - \lambda$

or, $\lambda' = \lambda + \Delta\lambda = 0.1 + 0.0242 = 0.1242 \text{ Å}$

Hence the frequency of the scattered photon

$$v' = \frac{c}{\lambda'} = \frac{3\times10^{8}}{0.1242\times10^{-10}} = 24 \times 10^{8} \times 10^{10} = 2.4 \times 10^{19} \text{ Hz.}$$

***Example 13.21.** An X-ray photon of wavelength 0.3 Å undergoes a 60° Compton scattering. Find the wavelength of the scattered photon and the kinetic energy imparted to the recoiling electron.*

Soln.

$$\Delta\lambda = \frac{h}{m_o c} = (1 - \cos\phi)$$

Here $\phi = 60°$, $\therefore$ $\cos 60° = 0.5$

$$\therefore \ \Delta\lambda = \frac{6.624\times10^{-34}}{9.1\times10^{-31}\times3\times10^{8}}(1 - 0.5)$$

$$= 0.0121 \times 10^{-10}\text{m}$$

$$= 0.0121 \text{ Å}.$$

Now $\Delta\lambda = \lambda' - \lambda$

or, $\lambda' = \lambda + \Delta\lambda = 0.3 + 0.0121 = 0.3121 \text{ Å}$

Kinetic energy of the recoil electron

$$\text{K.E.} = kv - hv'$$

$$= h\left[\frac{c}{\lambda} - \frac{c}{\lambda'}\right]$$

$$= h.c.\left[\frac{1}{\lambda} - \frac{1}{\lambda'}\right]$$

$$= hc\left[\frac{\lambda' - \lambda}{\lambda\lambda'}\right]$$

$$= 6.624 \times 10^{-34} \times 3 \times 10^{8}\left[\frac{0.0121\times10^{-10}}{0.3\times0.3121\times10^{-10}\times10^{-10}}\right]$$

$$\simeq 20 \times 10^{-26}\left[\frac{0.0121\times10^{-10}}{0.3\times0.3121\times10^{-10}\times10^{-10}}\right]$$

$$\simeq 20 \times 10^{-26} \times 0.13 \times 10^{10} \simeq 2.6 \times 10^{-16}\ \text{J}$$

$$= \frac{2.6\times10^{-16}}{1.6\times10^{-19}}\ \text{eV}$$

$$= 1.625 \times 10^{3}\ \text{eV}.$$

***Example 13.22.** In Compton scattering the angle of scattering was 30° while the angle of recoil was 60°. Calculate the energy of the incident photon.*

Soln.

we have

$$\tan \phi = \frac{\cot\left[\frac{\theta}{2}\right]}{1+\alpha} \quad \text{where } \alpha = \frac{h\nu}{m_o c^2}$$

Here $\phi = 60°$, and $\theta = 30°$

or, $\tan 60° = 1.732$ and $\cot\left[\frac{30}{2}\right] = \cot 15° = 3.732.$

$$\therefore \quad 1.732 = \frac{3.732}{1+\alpha}$$

or, $1 + \alpha = \frac{3.732}{1.732} = 2.154$

or, $\alpha = 1.154$

But $\alpha = h\nu/m_oc^2$

Hence, energy of the incident photon

$$h\nu = \alpha . m_oc^2$$

$= 1.154 \times 9.1 \times 10^{-31} \times (3 \times 10^8)^2$

$= 9.45 \times 10^{-14}$ J

$$= \frac{9.45\times10^{-14}}{1.6\times10^{-19}} \text{ eV}$$

$= 5.9 \times 10^5$ eV.

Example 13.23. *A photon of energy 5.1×10^5 eV is incident on aluminium foil. The photon is scattered at an angle of 90°. Calculate*

(i) wavelength of scattered photon

(ii) energy of recoil electron

(iii) direction of the emitted electron

Soln.

(i) $E = 5.1 \times 10^5 \text{eV} = 5.1 \times 10^5 \times 1.6 \times 10^{-19}$ J

$= 8.16 \times 10^{-14}$ J

$$E = h\nu = h.\frac{c}{\lambda}$$

or, $$\lambda = \frac{hc}{E} = \frac{6.624\times10^{-34}\times3\times10^8}{8.16\times10^{-14}}$$

$= 2.435 \times 10^{-12}$ m.

Now $$\Delta\lambda = \lambda' - \lambda = \frac{h}{m_o c}(1 - \cos\phi)$$

Here $\phi = 90°$ or, $\cos 90° = 0$

$$\therefore \quad \lambda' - \lambda = \frac{h}{m_o c} = \frac{6.624\times10^{-34}}{9.1\times10^{-31}\times3\times10^8}$$

$= 2.42 \times 10^{-12}$m.

$\therefore \quad \lambda' = 2.42 \times 10^{-12} + \lambda$

$= 2.42 \times 10^{-12} + 2.435 \times 10^{-12}$

$= 4.855 \times 10^{-12}$ m

$= 0.04855 \text{ Å}$

(ii) Energy of the recoil electron

$$= h\nu - h\nu'$$

$$= hc\left[\frac{1}{\lambda} - \frac{1}{\lambda'}\right]$$

$$= hc\left[\frac{\lambda' - \lambda}{\lambda\,\lambda'}\right]$$

$$= hc\left[\frac{\Delta\lambda}{\lambda\lambda'}\right]$$

$$= \frac{6.624\times10^{-34}\times3\times10^{8}\times2.42\times10^{-12}}{2.435\times10^{-12}\times4.855\times10^{-12}}$$

$$= 4.068 \times 10^{-14}\ \text{J}$$

$$= \frac{4.068\times10^{-14}}{1.6\times10^{-19}}\ \text{eV}$$

$$= 2.54 \times 10^{5}\ \text{eV}.$$

(iii) Direction of the recoil electron

$$\tan\varphi = \frac{\cot\left[\frac{\theta}{2}\right]}{1+\left[\frac{h}{m_o c\lambda}\right]} = \frac{\cot 45^\circ}{1+\frac{h}{m_o c\lambda}}$$

$$= \frac{1}{1+\left[\frac{6.624\times10^{-34}}{9.1\times10^{-31}\times3\times10^{8}\times2.435\times10^{-12}}\right]}$$

$$= 0.50$$

$$\therefore\quad \varphi = \tan^{-1}(0.50) = 26.6^\circ.$$

EXERCISE

1. Give an account of the production of X-rays. Discuss the properties and applications of X-rays.
2. What is the origin of X-rays? Distinguish between continuous X-radiations and characteristic X-ray emission spectra.
3. Describe Moseley's work on X-rays. What is Moseley's law? What is its importance?
4. Describe Laue's experiment and show how it established the wave nature of X-rays.
5. What is meant by absorption co-efficient of X-rays? Obtain an expression for the intensity of X-rays as it passes through an absorbing medium. What is half-value layer?
6. Derive Bragg's law for X-ray diffraction in crystals. How is it verified?
7. Describe and explain the X-ray spectrometer method of determining wavelength of X-rays.
8. Explain how X-ray spectrometer may be used to study the structure of crystals.
9. Describe the powdered crystal method of studying crystal structure.
10. What is Compton effect? Give the theory of Compton effect and obtain an expression for the change in wavelength of a photon undergoing Compton scattering. What is Compton wavelength?
11. Obtain expressions for the Kinetic energy and direction of the recoil electron in Compton scattering.
12. An X-ray tube operates at 12 KV. Find the maximum speed of the electrons striking the anticathode. Mass of electron = 9.1×10^{-31} kg; charge of electron = 1.602×10^{-19} C. **[6.5×10^7 ms^{-1}]**
13. NaCl has its principal planes spaced at 2.820 A.U. The first order of Bragg reflection is located at 10°. Calculate (a) the wavelength of the X-rays and (b) the angle for the second order Bragg reflection. **[0.9790 A.U.; 20°19']**
14. An X-ray tube is operated at 0.04 Megavolt. The radiations from the tube are analysed by a Bragg spectrometer. Calculate (a) shortest wavelength limit of the spectrum and (b) the glancing angle for the second order of this wavelength, given the crystal grating spacing = 3.036 A.U. **[0.309 A.U.; 2°55']**

15. The K_α line from molybdenum has a wavelength of 0.7078 A.U. Calculate the wavelength of the K_α line of copper. Atomic number of molybdenum = 42. Atomic number of copper = 29. [From Moseley's law, $(Z-1)^2 \lambda = \text{const}$. Let λ' be the wavelength of K_α line of copper. Then, $(42-1)^2 \times 0.7078 = (29-1)^2 \lambda'$ $\quad \therefore \lambda' = 1.517$ **A.U.]**
16. Monochromatic X-radiation of wavelength 0.124 A.U., undergoes Compton scattering from a carbon block. Calculate the wavelength that is scattered through 180°. **[0.1724 A.U.]**
17. Photon of energy 1.02 MeV undergoes Compton scattering through 180°. Calculate the energy of the scattered photon.

$$[\lambda = \frac{hc}{E} = \frac{6.624\times10^{-34}\times3\times10^{8}}{1.02\times1.6\times10^{-13}} = 0.012 \text{ A.U.}$$

$$d\lambda = 0.048 \text{ A.U.}; \lambda' = \lambda + d\lambda = (0.012 + 0.048) \text{ A.U.} = 0.060 \text{ A.U.}$$

$$\text{Energy of the scattered photon} = \frac{hc}{E} = 0.204 \text{ MeV}]$$

18. Explain why Compton effect is experimentally not observed for visible light rays. **[Maximum value of $\Delta\lambda = 0.048$ Å. Thus Compton effect can be detected only for those radiations whose wavelength is not greater than a few Å. For visible light ($\lambda \simeq 5000$ Å, $(\Delta\lambda)_{max}$ is only 0.001% of the initial wavelength which cannot be detected].**

CHAPTER XIV

The Atom Models

The first real foundation of the modern conception of atom was laid by Faraday who discovered that in electrolysis each atom, irrespective of the nature of the element, gave up or received a fixed quantity of positive or negative charge equal in magnitude to 1.59×10^{-19} coulomb. However, a more definite idea of the intrinsic nature of the atom came into existence with J.J. Thomson's discovery of the electron and his measurement of its charge to mass ratio e/m. As a result of this discovery the following two facts were clearly established – (a) electrons enter into the constitution of all atoms and (b) since the atom as a whole is electrically neutral the quantity of positive and negative charges in it must be same.

14.1 The Thomson Atom Model

With the above two conclusions to guide him, J.J. Thomson in 1898 proposed that the positive charges of an atom is uniformly distributed in a sphere of atomic dimension while the electrons were so arranged inside the positive sphere that their mutual repulsions were exactly balanced by the force of attraction towards the centre of the sphere. Thomson was able to detail the arrangement of electrons ranging from 1 to 100 inside the positive sphere. Since the electrons are distributed throughout the atom like raisins in a plum pudding, this model is sometimes known as *plum-pudding* model.

The force on an electron at a distance r from the centre of a uniformly charged sphere of radius R can be computed using Gauss' law.

$$F = \frac{Ze^2}{4\pi\epsilon_0 R^3} r = kr \tag{14.1}$$

As a result of this linear restoring force, the electrons would be able to oscillate about their equilibrium position just like a mass on a spring subject to the linear restoring force F = kx.

Thus the electrons in the Thomson model are expected to oscillate about their equilibrium positions with a frequency $\nu = (2\pi)^{-1}\sqrt{k/m}$, where k is the constant defined by eqn. (14.1). Since an oscillating electric charge radiates electromagnetic waves whose frequency is identical to the oscillation frequency, it was expected that the radiation emitted by atoms would show this characteristic frequency. This turns out not to be true. Considering the case of the simplest hydrogen atom, it was found that the above assumed mechanism should give rise to a spectral line in the whereabouts of $1400\,\text{Å}$. This corresponds roughly to a line in the observed hydrogen spectrum. However, hydrogen gives rise not only to a single spectral line but to several series, each series consisting of several lines.

The most serious failure of the Thomson atom model arises from the scattering of charged particles by atoms. Let us consider the passage of a single positively charged particle through the atom. In course of its passage through the atom, the particle is deflected somewhat from its original trajectory owing to the electrical forces exerted on the particle by the atom. These forces are (i) a repulsive force due to the positive charge of the atom, and (ii) an attractive force due to the negatively charged electrons. Let us assume that the mass of the deflected particle is both much greater than the mass of an electron and yet much less than the mass of the atom. In the encounter between the projectile and an electron, the forces exerted on each other are, according to Newton's third law, equal and opposite. Hence the principal victim of the encounter is the much less massive electron; the effect on the projectile will be negligible. Thus only the positively charged atom needs to be considered as a cause of the deflection of the projectile. By the same argument, any possible motion of the more massive atom caused by the passage of the projectile can be neglected. The experiment therefore reduces to the scattering of a positively charged projectile by the stationary positively charged massive part of the atom.

Fig. 14.1 shows a schematic representation of the deflection of the projectile, which is moving with a speed v along a line that would pass a distance b from the centre of the atom if the projectile were not deflected. The distance b is called the *impact parameter*. It will be assumed that $v \ll c$ so that non-relativistic mechanics could

be used and so, $K = \frac{1}{2}mv^2$. The electrical repulsion causes a small deflection and so the particle leaves the atom moving in a slightly different direction, at an angle θ with respect to the original direction.

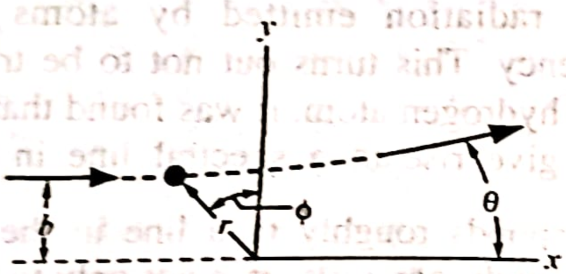


Fig. 14.1

The angle of deflection θ can be calculated by considering the *impulse* received by the projectile, which gives it some momentum in the y-direction.

$$\Delta p_y = \int F_y \, dt \qquad (14.2)$$

At an arbitrary point along the trajectory,

$$F_y = F \cos \phi$$

The projectile, which has a charge $q = ze$, experiences a force F given by analogy with eqn. (14.1):

$$F = \frac{1}{4\pi \epsilon_0} \frac{z Ze^2}{R^3} r = zkr \qquad (14.3)$$

where k is the same constant as defined by eqn. (14.1). Since $\cos \phi \cong b/r$, we have

$$\Delta p_y \cong \int zkr.\frac{b}{r}.dt = zkb \int dt$$

$$= zkbT \qquad (14.4)$$

where T is the total time taken by the projectile to travel through the atom, which is the total distance traveled in the atom divided by the average speed. Since the deflection is small, the path can be approximated as a straight line, as shown in Fig. 14.2 and the average speed is very nearly equal to v. So

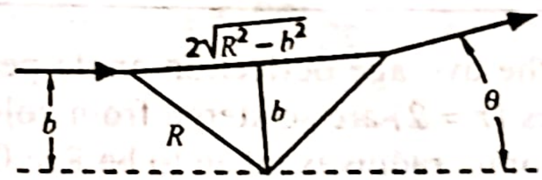


Fig. 14.2

$$T \cong \frac{2\sqrt{R^2-b^2}}{v} \tag{14.5}$$

Eqn. (14.4) then gives

$$\Delta p_y \cong \frac{2zkb}{v}\sqrt{R^2-b^2} \tag{14.6}$$

Assuming that the change in p_x is negligible, we have

$$\tan\theta = \frac{p_y}{p_x} \cong \frac{\Delta p_y}{p} \tag{14.7}$$

If θ is small, $\tan\theta \cong \theta$, so

$$\theta \cong \frac{\Delta p_y}{p} = \frac{2zkb}{mv^2}\sqrt{R^2-b^2} \tag{14.8}$$

The value of the scattering angle θ depends on the value of the impact parameter b. When b = 0, the projectile is undeflected (θ = 0) because the net force on the projectile is zero. When b = R, the projectile is again undeflected, because it is assumed that the force acts only when the projectile is within the atom. A representative

average value of θ can be found corresponding to an average value of b, for which the intermediate value b = R/2 is chosen.

In that case eqn. (14.8) gives

$$\theta_{avg} = \sqrt{\frac{3}{4}\frac{zkR^2}{mv^2}} \tag{14.9}$$

For example the average deflection angle per collision when 5 MeV alpha particles (z = 2) are scattered from gold (Z = 79) is given below. A typical atomic radius is taken to be R = 0.1 nm.

The quantity zkR^2 can be computed to be

$$zkR^2 = 2\left(\frac{Ze^2}{4\pi\epsilon_0 R^3}\right)R^2 = (2)(79)\frac{1.44\text{eV.nm}}{0.1\text{nm}} \cong 2.3\text{keV}$$

$$\left[\frac{e^2}{4\pi\epsilon_0} = 1.44\text{ eV.nm}\,;\ \text{Also } mv^2 = 2K = 10\text{ MeV}\right]$$

$$\text{so, } \theta_{avg} = \sqrt{\frac{3}{4}\frac{2.3\text{keV}}{10\text{MeV}}} \cong 2\times10^{-4}\text{rad} = 0.01^0$$

When a beam of projectile is incident on a thin foil, instead of the scattering of a single projectile by a single atom, what is observed is the scattering of many projectiles by many atoms. One does not have any control over the impact parameter b for each scattering. In passing through a foil, a projectile is scattered many times. For a typical foil thickness of 1 μm (10^{-6} m), the projectile is scattered by about 10^4 atoms, each of which deflects the projectile by an angle whose average value is θ_{avg}. The total scattering angle for any particular projectile is determined by statistical considerations, since some of the individual scatterings move the projectile toward larger scattering angles and some toward smaller angles. For N scatterings, the average net scattering angle θ is related to the average individual scattering angle by

$$\theta \simeq \sqrt{N}\,\theta_{avg} \tag{14.10}$$

For $N = 10^4$ and $\theta_{avg} = 0.01°$, the average net scattering angle is expected to be about 1°, which is consistent with observations.

The Thomson model, however, fails when the probability of large angle scattering is examined. If each individual scattering deflects the projectile through an angle of 0.01°, then to observe projectiles scattered through a total angle greater than 90°, one must have about 10^4 successive scatterings, *all* of which push the projectile toward larger angles. Since the probabilities of individual scatterings toward either larger or smaller angles are equal, the probability of having 10^4 successive scatterings toward larger angles, like the probability of finding 10^4 successive heads in tossing a coin, is about $\left(\frac{1}{2}\right)^{10{,}000} = 10^{-3{,}000}$.

Hans Geiger and Ernest Marsden performed an experiment of this sort in the laboratory of Ernest Rutherford in 1910. Their result showed that the probability of an alpha particle scattering at angles greater than 90° was about 10^{-4}. This remarkable discrepancy between the expected value (10^{-3000}) and the observed value (10^{-4}) was described by Rutherford in this way:

It was quite the most incredible event that ever happened to me in my life. It was as incredible as if you fired a 15 inch shell at a piece of tissue paper and it came back and hit you.

14.2 Rutherford Nuclear Atom Model

In analyzing the scattering of alpha particles, Rutherford concluded that the most likely way an alpha particle (m = 4u) can be deflected through large angles is by a single collision with a more massive object. He therefore proposed, in 1911, that the mass and positive charge of the atom were not distributed uniformly over the volume of the atom, but instead were concentrated in an extremely small region, about 10^{-14}m in diameter, at the centre of the atom. This central core of the atom in which the entire positive charge of the atom and almost the entire mass of the atom was concentrated was called the *nucleus*. The electrons in the atom were, in consequence, assumed to be situated outside the nucleus in some

sort of configuration. Thus Rutherford arrived at the concept of nuclear atom model.

Applying the same laws of probability to this atom model it was shown that the proportion of alpha particles deflected through large angles was much greater than in the case of Thomson atom model. A projectile of charge ze, experiences a repulsive force due to the positively charged nucleus:

$$F = \frac{(ze)(Ze)}{4\pi\epsilon_0 r^2}$$

It is assumed that the projectile is always outside the nucleus, so that it feels the full nuclear charge Ze. The atomic electrons, with their small mass, do not appreciably affect the path of the projectile and their effect on the scattering is neglected. It is further assumed that the nucleus is so much more massive than the projectile that it does not move during the scattering process; since no recoil motion is given to the nucleus, the initial and final kinetic energies K of the projectile are equal.

Theory of alpha particle scattering

In order to find a theory of alpha particle scattering by thin metal foils, Rutherford made the following assumptions:

(i) The alpha particle and the nucleus with which it interacts are small enough to be considered as point masses and charges.

(ii) The only force acting between the alpha particle and the nucleus is the electrostatic repulsive force between them.

(iii) Compared to the alpha particle the nucleus is so massive that its motion during the impact may be disregarded.

The scattering of alpha particle by the nucleus is treated as the elastic impact of two particles. Owing to the variation of the electrostatic force with $1/r^2$, where r is the instantaneous distance between the alpha particle and the nucleus, the electrostatic force of repulsion between them would continue to increase enormously as the alpha particle gets closer and closer to the nucleus.

By considering the motion under central forces, it can be shown that the path of the alpha particle will change from straight line to a hyperbola PAP' with the nucleus N at the outer focus (Fig. 3). The angle between asymptotic direction PO in which the alpha particle approaches the nucleus and the asymptotic direction OP' in which it

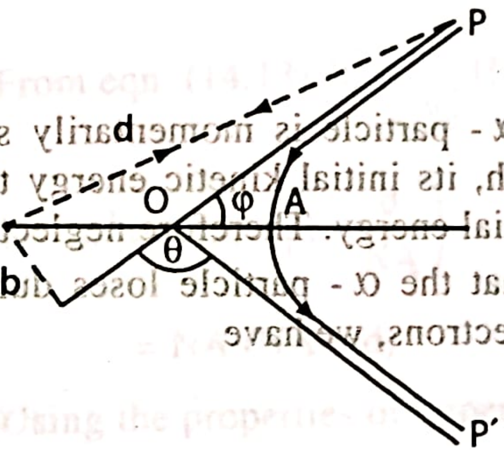


Fig. 14.3(a)

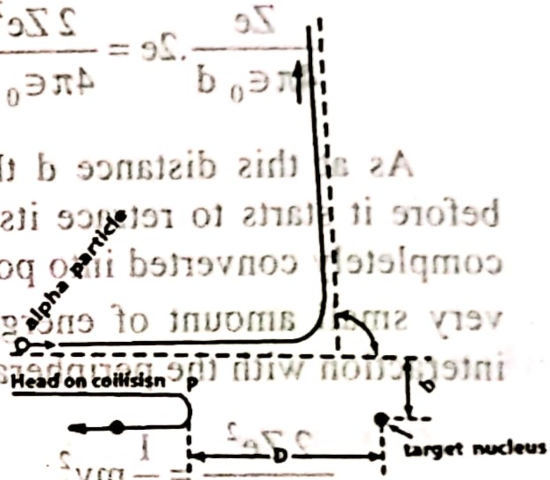


Fig. 14.3(b)

recedes, is the scattering angle θ. Let the minimum distance to which the alpha particle would approach the nucleus in the absence of any force between them be b. This is the perpendicular distance of N from OP and is known as the *impact parameter*. Let m be the mass and v_0 be the initial velocity of the alpha particle. If Z be the atomic number of the element which is being used as scatterer of the α - particle, then the charge on the nucleus would be Ze.

To find a relation between b and θ, let us consider the case of a head on collision *i.e.*, when the α - particle is directed straight towards the nucleus N so that b = 0. As the α - particle gets closer and closer to the nucleus, its kinetic energy will get less and less due to the strong repulsive force between them until at a certain distance 'd' from the nucleus the α - particle would be forced to retrace its path, in which case the angle of scattering θ would be 180°. This distance d is the distance of closest approach between the α - particle and the nucleus and should not be confused with the impact parameter b. The electrostatic potential at a distance d from the

nucleus is given by $\frac{Ze}{4\pi\epsilon_0 d}$ and this acts on the α - particle of charge ze. Hence the potential energy of the α - particle when it is at a distance d from the nucleus is

$$\frac{Ze}{4\pi\epsilon_0 d}.2e = \frac{2Ze^2}{4\pi\epsilon_0 d}$$

As at this distance d the α - particle is momentarily stopped before it starts to retrace its path, its initial kinetic energy there is completely converted into potential energy. Therefore neglecting the very small amount of energy that the α - particle loses during its interaction with the peripheral electrons, we have

$$\frac{2Ze^2}{4\pi\epsilon_0 d} = \frac{1}{2}mv_0^2$$

$$d = \frac{4Ze^2}{4\pi\epsilon_0 mv_0^2} \tag{14.11}$$

In practice it is not possible to direct the α - particle exactly towards the nucleus. We will therefore have to consider when $b \neq 0$, and the α - particle will be scattered through an angle θ which is less than 180° and travel along the hyperbolic path PAP'.

Let v be the velocity of the α - particle at the vertex A. Applying the principle of conservation of energy and momentum, we have

$$\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 + \frac{2Ze^2}{4\pi\epsilon_0 NA} \tag{14.12}$$

$$mv_0b = mv.\ NA \tag{14.13}$$

where NA is the distance of the vertex A from the nucleus N.

Substituting the value of Ze^2 from eqn. (14.11) in eqn. (14.12) we get

$$\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 + \frac{mv_0^2 d}{2NA}$$

$$\therefore \quad v^2 = v_0^2\left(1-\frac{d}{NA}\right)$$

$$\text{or,} \quad \frac{v^2}{v_0^2} = \left(1-\frac{d}{NA}\right)$$

From eqn. (14.13) $b^2 = \frac{v^2}{v_0^2}(NA)^2$

$$\therefore \quad b^2 = (NA)^2\left(1-\frac{d}{NA}\right)$$

$$= NA\,(NA - d) \qquad (14.14)$$

Using the properties of hyperbola, namely the eccentricity,

$$\in = \frac{1}{\cos\phi}, \text{ where } \phi = \frac{1}{2}(\pi - \theta) \text{ and } NO = \in .OA,$$

we have

$$NA = NO + OA = NO\left(1+\frac{1}{\in}\right) = NO\,(1 + \cos\phi)$$

As $NO = \frac{b}{\sin\phi}$

$$\therefore \quad NA = \frac{b}{\sin\phi}(1 + \cos\phi)$$

$$= b\left(\frac{1+2\cos^2\frac{\varphi}{2}-1}{2\sin\frac{\varphi}{2}\cos\frac{\varphi}{2}}\right)$$

$$= \frac{b\cos\frac{\varphi}{2}}{\sin\frac{\varphi}{2}} = b\cot\frac{\varphi}{2}$$

Putting this in eqn. (14.14)

$$b^2 = b\cot\frac{\varphi}{2}\left(b\cot\frac{\varphi}{2} - d\right)$$

$$d = b.\frac{\cot^2\frac{\varphi}{2} - 1}{\cot\frac{\varphi}{2}} = 2b\cot\varphi$$

$$= 2b\cot\left(\frac{\pi-\theta}{2}\right) \qquad \because \varphi = \frac{\pi-\theta}{2}$$

or, $$d = 2b\tan\frac{\theta}{2}$$

Substituting the value of d from (14.11) we get

$$\tan\frac{\theta}{2} = \frac{d}{2b} = \frac{Ze^2}{2\pi\epsilon_0\, bmv_0^2} \qquad (14.15)$$

It can be seen from eqn. (14.15) that for fixed values of Z, m and v_o the scattering angle θ increases from 0 to 180° as the impact parameter b decreases from relatively high values to the limiting value zero. It can be concluded that if b is large and the α - particle passes far away from the nucleus the scattering angle is very small. In the limiting case when the α - particle is directed straight towards the nucleus (b = 0), θ = 180° and the α - particle will retrace its path after approaching the nucleus upto a distance d. The scattering of α - particle for different values of b is shown in Fig. 14.4.

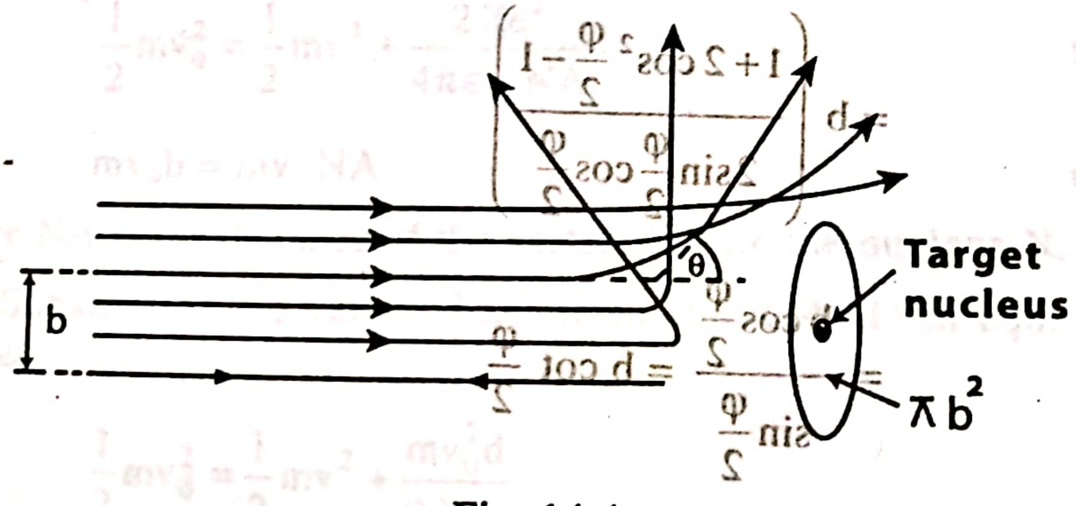


Fig. 14.4

It is also clear from eqn. (14.15) that all the α - particles approaching a target nucleus with an impact parameter from 0 to b will be scattered through an angle θ or more. This means that an α - particle, initially directed anywhere within the area πb^2 around the nucleus, will be scattered through an angle θ or more (Fig. 14.4). The area πb^2 is accordingly called the *cross section* for interaction and is generally denoted by σ.

It should however be pointed out here that the incident α - particle does not necessarily pass within a distance b of the nucleus, because it is actually scattered before it can reach the vicinity of the nucleus.

We shall now deal with the situation which can be realized in an actual experiment. Let a narrow beam of α - particle be directed normally towards the target foil (Fig. 14.5). Let the thickness of the foil be t and let the number of atoms per unit volume in it be n. The number of target nuclei per unit area will therefore be nt and a beam of α - particle incident upon an area A will encounter ntA nuclei. As the cross section per nucleus for an α - particle to be scattered through an angle θ or more is σ, the aggregate cross section for scattering through θ or more by ntA nuclei will be ntAσ. Hence the fraction f of the incident α - particles that are scattered through an angle θ or more is given by

$$f = \frac{\alpha\text{ - particles scattered through } \theta \text{ or more}}{\text{incident } \alpha \text{ - particles}}$$

$$= \frac{\text{Aggregate cross section}}{\text{target area}}$$

$$= \frac{ntA\sigma}{A} = nt\sigma = nt\pi b^2.$$

Putting the value of b from eqn. (14.15), we have

$$f = nt\,\pi \left(\frac{2Ze^2}{4\pi\epsilon_0 \, mv_0^2} \right)^2 \cot^2 \frac{\theta}{2} \qquad (14.16)$$

Although a real scattering foil may be thousands or tens of thousands of atoms thick, eqn. (14.16) was derived on the assumption that the foil was one atom thick – a single layer of atoms packed tightly together, so that there is no overlapping of the cross section of adjacent nuclei and a scattered α - particle receives its deflection from a single encounter with a nucleus.

An experimental arrangement for measuring the number of α - particles scattered though an angle θ and $\theta + d\theta$ is shown in Fig.14.5. The number of such α - particles can be obtained by differentiating eqn. (14.16) with respect to θ.

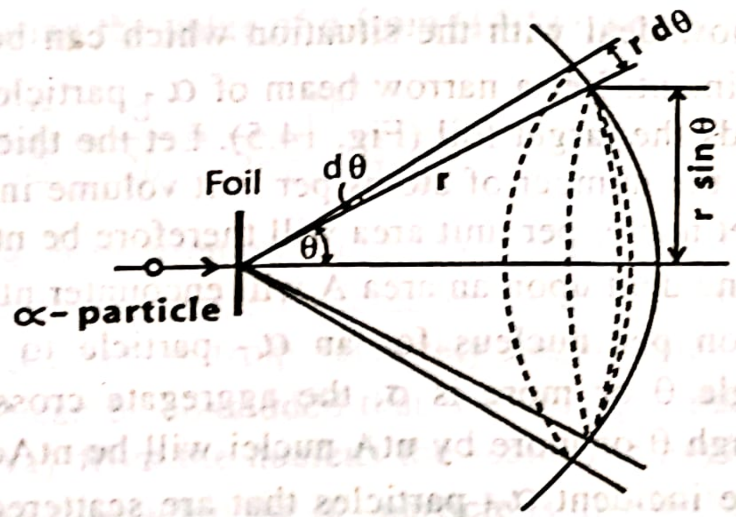


Fig. 14.5

$$df = -\, nt\pi \left(\frac{Ze^2}{2\pi \epsilon_0 \, mv_0^2} \right)^2 \cot \frac{\theta}{2} \, \text{cosec}^2 \frac{\theta}{2} \, d\theta \qquad (14.17)$$

Eqn. (14.17) represents the fraction of the incident particles scattered between θ and $\theta + d\theta$; the negative sign indicating that the fraction decreases with increasing θ.

In the experiment, a fluorescent screen is placed behind the foil, say, at a distance r and the α - particles scattered between θ and $\theta + d\theta$ strike a zone of sphere of radius r whose width is $rd\theta$ and radius $r \sin\theta$. The area of the screen struck by these particles is therefore given by

$$dS = \text{circumference} \times \text{width}$$

$$= 2\pi r \sin\theta \times r d\theta$$

$$= 4\pi r^2 \sin\frac{\theta}{2} \cos\frac{\theta}{2} d\theta.$$

Let the total number of α - particles that strike the foil during the course of the experiment be N_i. The number of α - particles scattered into an angle $d\theta$ at θ would then be $N_i df$. If $N(\theta)$ be the number of α - particles actually measured in the experiment, then

$$N(\theta) = \frac{N_i \, df}{dS}$$

$$= \frac{N_i \, n\pi t \left(\frac{Ze^2}{4\pi \epsilon_0 m v_0^2} \right)^2 \cot\frac{\theta}{2} \operatorname{cosec}^2\frac{\theta}{2} d\theta}{4\pi r^2 \sin\frac{\theta}{2} \cos\frac{\theta}{2} d\theta}$$

$$= \frac{Z^2 e^4 n t N_i}{64\pi^2 \epsilon_0^2 m^2 v_0^4 r^2 \sin^4\frac{\theta}{2}} \qquad (14.18)$$

Eqn. (14.18) represents the Rutherford scattering formula and indicates that the number of α - particles per unit area reaching the screen a distant r from the foil is

(i) directly proportional to the thickness of the foil (t)

(ii) directly proportional to the square of the atomic number of the foil (Z^2)

(iii) inversely proportional to $\sin^4\frac{\theta}{2}$

and (iv) inversely proportional to the square of the initial kinetic energy $(m v_0^2)^2$ or to the fourth power of the initial velocity (v_0^4) of the α - particles.

Experimental verification of Rutherford Scattering

In 1913, the theoretical predictions of Rutherford scattering formula were tested and verified in Rutherford's laboratory by Geiger and Marsden in a remarkable series of experiments requiring great care and experimental skill. The schematic diagram of the experimental arrangement is given in Fig. 14.6. A radioactive source of alpha - particles is in a shield with small hole. α - particles strike the foil F and are scattered into the angular range dθ. In those days before electronic recording and processing equipment was available, Geiger and

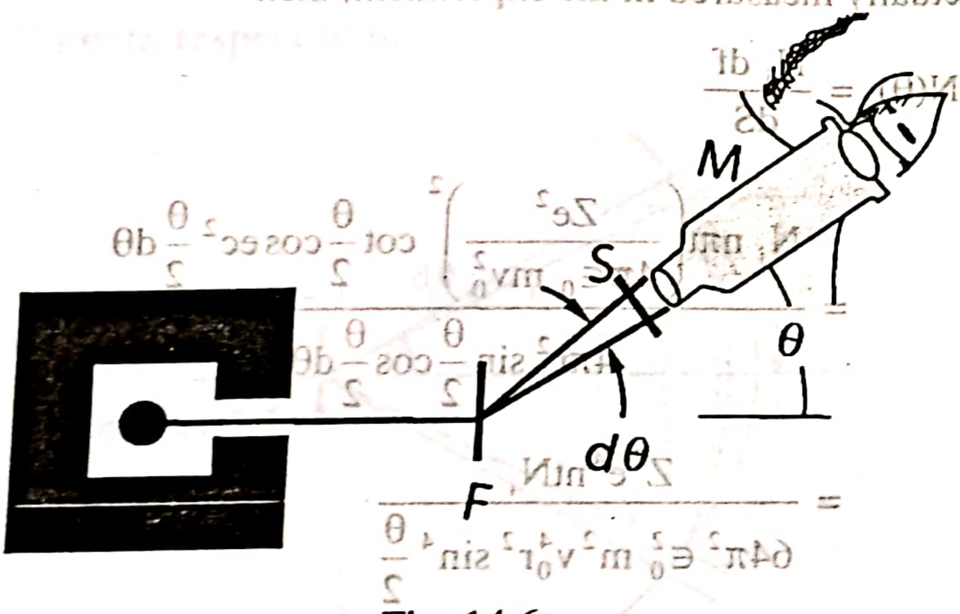


Fig. 14.6

Marsden observed and recorded the alpha particles by counting the scintillations (flashes of light) produced when the alpha particles struck a zinc sulphide screen. Each time a scattered particle strikes the screen S a flash of light is emitted and observed with the movable microscope M. In all four predictions of the Rutherford scattering formula were tested:

(i) $N(\theta) \propto t$: With a source of 8-MeV alpha particles from radioactive decay, Geiger and Marsden used scattering foils of varying thicknesses t while keeping the scattering angle θ fixed at 25°. Their results are summarized in Fig. 14.7. The linear dependence of N (θ) on t is quite apparent. What is even more significant is the evidence that even at this moderate scattering angle, *single* scattering is much more important than *multiple* scattering.

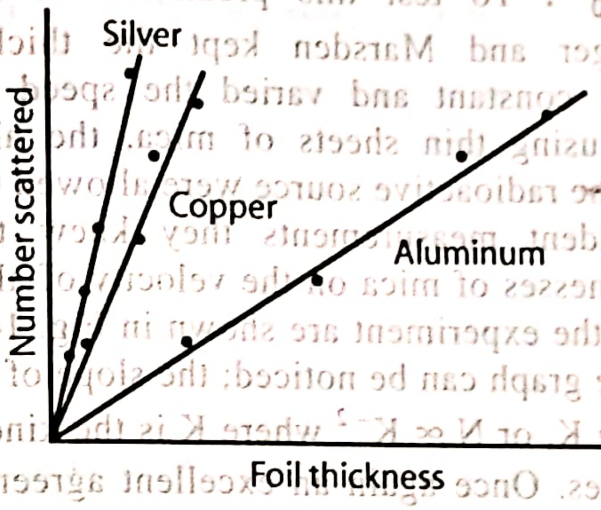


Fig. 14.7

(ii) $N(\theta) \propto Z^2$: Geiger and Marsden used a variety of different scattering materials for this part of the experiment. The materials were of approximately (but not exactly) the same thickness. Consequently this part of the experiment was much more difficult to test than the previous one, since it involved the comparison of *different thicknesses of different materials.* However, as shown in Fig. 14.8, the result is consistent with the proportionality of $N(\theta)$ to Z^2.

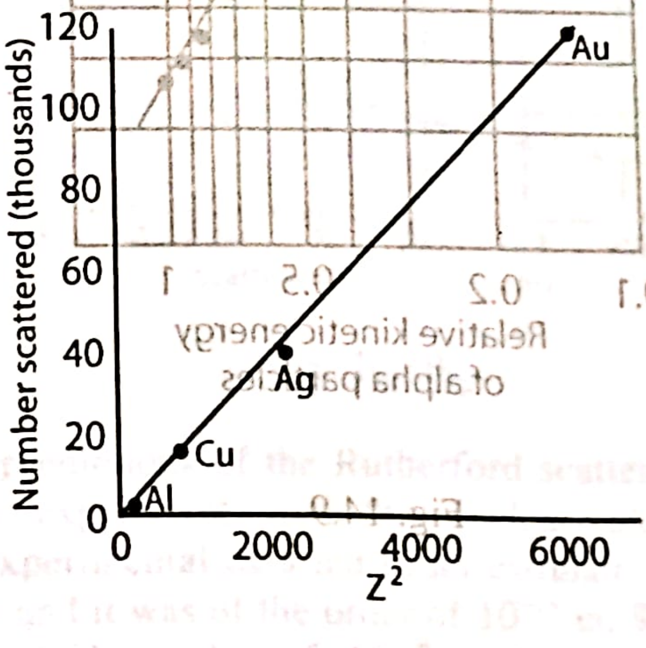


Fig. 14.8

(iii) $N(\theta) \propto 1/mv_0^2$: To test this prediction of the scattering formula, Geiger and Marsden kept the thickness of the scattering foil constant and varied the speed of the alpha particles. By using thin sheets of mica, the alpha particles emitted from the radioactive source were allowed to slow down. From independent measurements they knew the effect of different thicknesses of mica on the velocity of alpha particles. The results of the experiment are shown in Fig. 14.9. The log - log scale of the graph can be noticed; the slope of -2 shows that log N $\propto$ –2 log K, or N $\propto$ K^{-2} where K is the kinetic energy of the α – particles. Once again an excellent agreement with the expected relationship is obtained.

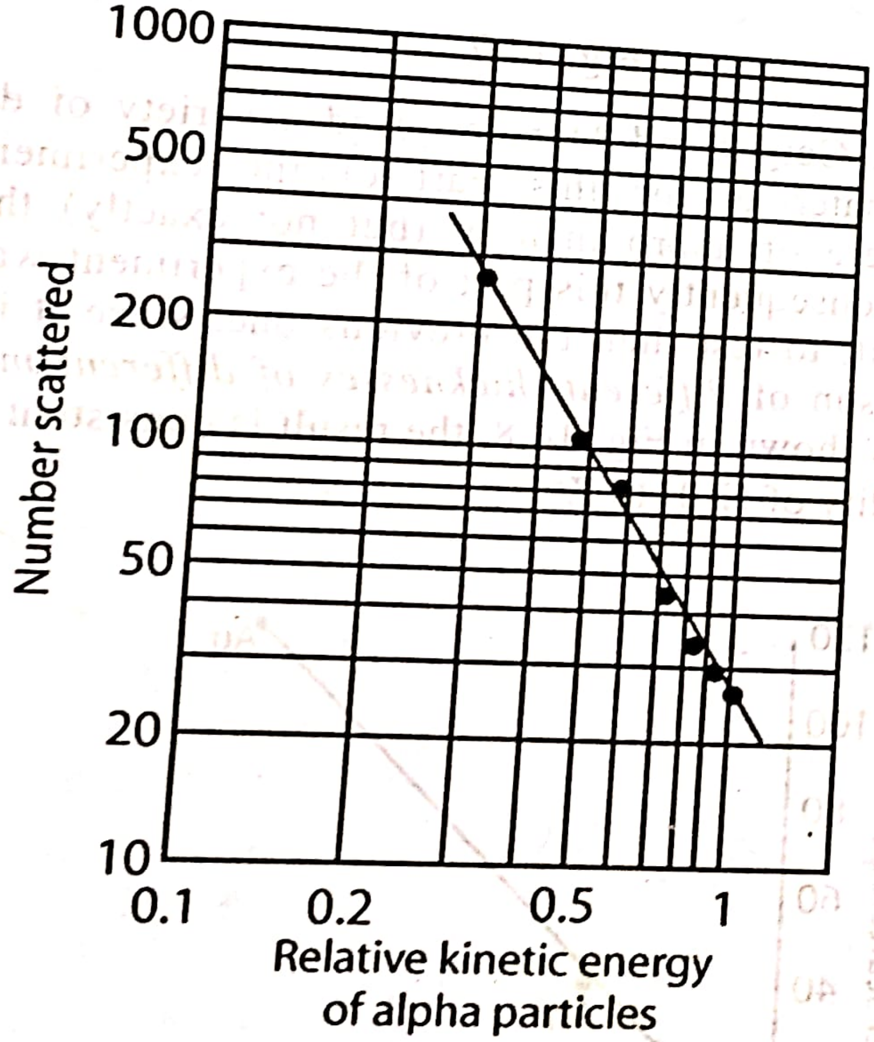


Fig. 14.9

(iv) $N(\theta) \propto \sin^{-4}\dfrac{\theta}{2}$: The dependence of N on θ is perhaps the most important and distinctive feature of the Rutherford scattering formula. It also produces the largest variation in N over the range accessible by experiment. In the previous tests N varies by perhaps an order of magnitude; in this case N varies by about *five* orders of magnitude from the smaller to the larger angles. Geiger and Marsden used a gold foil and varied θ from 5 to 150°, to obtain the relationship between N and θ. The results are shown in Fig. 14.10. The $\sin^{-4}\left(\dfrac{\theta}{2}\right)$ dependence is exactly as predicted by the Rutherford formula.

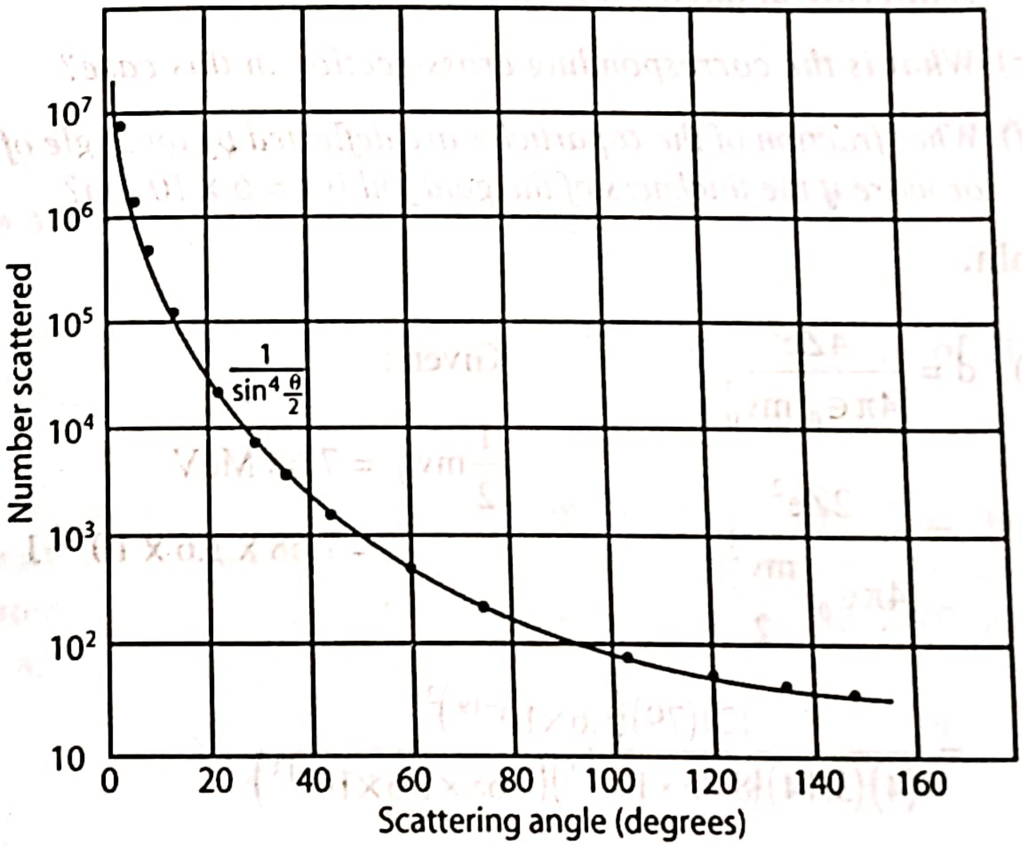


Fig. 14.10

Thus all predictions of the Rutherford scattering formula were confirmed by experiment, and the *nuclear atom* was verified. Further, the experimental data led to an estimate of the dimensions of the nucleus and it was of the order of 10^{-14} m. Since dimension of the atom is of the order of 10^{-10}m, the nucleus occupies an extremely small part of the atom. In the space between the nucleus

and the outer confines of the atom there are a few electrons, whose masses are very small and dimensions more or less of the same order of magnitude as that of the nucleus, but whose charges exactly balance the positive charges concentrated in the nucleus. Hence it is normally said that the atom consists chiefly of empty space.

***Example 14.1** Geiger and Marsden used α-particles of kinetic energy 7.68 MeV and a gold foil in their experiment. Given that for gold Z = 79.*

(a) Calculate the distance d of closest approach to a gold nucleus.

(b) Calculate the impact parameter b in order to produce scattering at angles θ ≥ 90°.

(c) What is the corresponding cross-section in this case?

(d) What fraction of the α-particles are deflected by an angle of 90° or more if the thickness of the gold foil is t = 6 × 10⁻⁷ m?

Soln.

(a)
$$d = \frac{4Ze^2}{4\pi\epsilon_0 mv_0^2}$$

Given :
$$\frac{1}{2}mv_0^2 = 7.68 \text{ MeV} = 7.68 \times 1.6 \times 10^{-13} \text{ J}$$

$$= \frac{2Ze^2}{4\pi\epsilon_0 \dfrac{mv_0^2}{2}}$$

$$= \frac{(2)(79)(1.6\times10^{-19})^2}{(4)(3.14)(8.85\times10^{-12})(7.68\times1.6\times10^{-13})}$$

$= 2.96 \times 10^{-14}$ m.

(b) $\tan\dfrac{\theta}{2} = \dfrac{d}{2b}$; or, $b = \dfrac{d}{2\tan\dfrac{\theta}{2}} = \dfrac{d}{2}\cot\dfrac{\theta}{2}$

$$\therefore b = \frac{2.96\times10^{-14}}{2}\cot\frac{90^\circ}{2}$$

$= 1.48 \times 10^{-14}$ m.

(c) $\sigma = \pi b^2$

$= (3.14)(1.48 \times 10^{-14})^2$

$= 6.87 \times 10^{-28}\ m^2$

$= 6.87$ barns.

(d) The number of nuclei per unit volume,

$$n = \frac{\rho N_A}{M}$$

where ρ = density of the material of the foil

$= 1.93 \times 10^4$ Kg/m^3 for gold

N_A = Avogadro's number = 6.02×10^{26} atoms/Kg-mole

and M = atomic weight of the material of the foil (Kg/Kg-mole)

= 197 (for gold)

$$\therefore \quad n = \frac{\rho N_A}{M} = \frac{(1.93\times10^4)(6.02\times10^{26})}{197} = 5.91 \times 10^{28} \text{ atoms/m}^3$$

Now $f = n\sigma t = (5.91 \times 10^{28})(6.87 \times 10^{-28})(6 \times 10^{-7})$

$= 2.44 \times 10^{-5}$

This means that only approximately two out of 100,000 particles suffer deflection of 90° or more.

Example 14.2 Assume that the charge +Ze of a nucleus is spread over a sphere of radius R. Then show that the fastest alpha-particles (charge 2e, mass 4m_p) which can suffer 180° scattering will have speed v given by

$$v = (e^2Z/m_pR)^{1/2}$$

With gold (Z = 79), 180° scattering is observed with alpha-particles of v upto 1.6 x 10^9 cm/sec. Deduce the upper limit for the radius of the nucleus.

Soln.

The nucleus being spherical, the charge Ze may be considered to be concentrated at its centre. An alpha-particle will have its smallest distance of closest approach d from the centre of the

nucleus when its impact parameter b = 0 corresponding to a head-on approach resulting in a 180° scattering. Now the distance of closest approach, even for the fastest α-particle, cannot be less than the radius R of the nucleus. Therefore for the fastest α-particle,

$$d = R$$

At the instant of closest approach, the α-particle momentarily stops at a distance R from the centre of the nucleus. At this moment, its kinetic energy is entirely converted into potential energy. Therefore, at that instant

$$\frac{2Ze^2}{R} = \frac{1}{2}mv^2$$

$$= \frac{1}{2}(4m_p)\,v^2$$

$$\therefore \quad v = \left(\frac{Ze^2}{m_p R}\right)^{1/2}$$

From this the upper limit for the radius of the sphere is

$$R = \frac{Ze^2}{m_p v^2}$$

Putting Z = 79, $v = 1.6 \times 10^9$ cm/sec,

$m_p = 1.67 \times 10^{-24}$ gm and $e = 4.8 \times 10^{-10}$ e.s.u.

$$R = \frac{(79)(4.8\times10^{-10})^2}{(1.67\times10^{-24})(1.6\times10^9)^2}$$

$$= 4.26 \times 10^{-12} \text{ cm.}$$

Drawbacks of Rutherford Nuclear Atom

Although his nuclear atom model was at once universally accepted, Rutherford himself soon realized that his model was not free from limitations, the chief among them arising from the

considerations of the distribution of the electrons outside the nucleus and the stability of the atom as a whole. For it became obvious that in the nuclear atom *equilibrium could not be secured by the operation of the electrostatic forces alone* between the positively charged nucleus and the negatively charged electrons outside the nucleus. For instance, let us consider the case of an atom with two electrons; its nuclear charge is +2e. If the electrons are symmetrically placed at a distance r from the nucleus, the force of attraction between the nucleus and each of the electrons is $\frac{2e^2}{4\pi\epsilon_0 r^2}$, while the force of repulsion between the electrons is $\frac{e^2}{4\pi\epsilon_0 (4r^2)}$.

Thus the force of attraction is eight times greater than the force of repulsion and hence the condition of stability is not satisfied. The electrons would fall into the nucleus, thus destroying the stable structure of the atom.

In order to overcome this difficulty, Rutherford suggested that the *electrons might be assumed to revolve round the nucleus*, like the planets round the sun, at such a speed that the mechanical centrifugal force would just balance the net excess of electrostatic attraction and in consequence the stability of the atom could be secured.

But such an assumption brought, in its wake, another very serious difficulty from the point of view of electromagnetic theory, according to which *a revolving electron should radiate energy continuously.* Now this energy can only come from the atomic system. The atom therefore would steadily lose energy. As a result, the electron will approach the nucleus by a spiral path, giving out a radiation of constantly increasing frequency and finally fall into the nucleus. Thus the orbital motion of the electrons destroys the very purpose for which it was postulated *viz.*, the stability of the atom. Moreover, emission of radiation of constantly increasing frequency has no experimental basis. Atoms are actually found to emit discrete spectral lines of definite frequency. Thus the conclusion is very clear : either the Rutherford nuclear atom model with revolving electrons is defective or the classical electromagnetic theory fails in this case. Niels Bohr in 1913 solved this dilemma by admitting the failure of the classical theory and applying with remarkable success the quantum theory to the Rutherford nuclear atom with revolving electrons.

14.3 The Hydrogen Spectrum

The radiation from atoms can be classified into continuous spectra and discrete or line spectra. In a continuous spectrum, all wavelengths from some minimum, perhaps 0, to some maximum, perhaps approaching ∞, are emitted. Incandescant solids and liquids, when white hot, emit continuous spectrum (*e.g.*, the solar spectrum). If, on the other hand, an electric discharge is forced in a tube containing a small amount of the gas or vapour of a certain element such as mercury, sodium, or neon, light is emitted at a few discrete wavelengths and not at any others. The blue-green tint of the mercury vapour street lights is due to the strong 436 nm (blue) and 546 nm (green) lines in the mercury spectrum. Similarly the softer, yellowish colour of the sodium vapour street lights is due to the strong yellow line at 590 nm in the sodium spectrum (which is actually a *doublet* – two very closely spaced lines). The intense red lines of neon are responsible for the red colour of "neon signs".

Another type of spectra is the absorption spectrum. When a beam of white light, containing all wavelengths, passes through a sample of gas, it will be found that certain wavelengths have been absorbed from the light, and again a line spectrum results. These wavelengths correspond to many (but not all) of the wavelengths seen in the emission spectrum.

Since the interpretation of line spectra is very difficult in complex atoms, the line spectra of the simplest atom, hydrogen will be discussed here. Regularities appear in both the emission and absorption spectra. As in the case with the mercury and sodium spectra, some lines present in the emission spectrum are missing from the absorption spectrum.

To explain the spectra of hydrogen atom, a "reverse scientific method" approach was used. As in the case of blackbody radiation, in which, in the absence of a theory that explains the data, first a function was found that fits the data, and then attempt was made to find a theory that explains the derived function. In 1885 Johannes Balmer, a Swiss school teacher, noticed that the wavelengths of the group of emission lines of hydrogen in the visible region could be very accurately calculated from the formula

$$\lambda = 364.5 \frac{n^2}{n^2 - 4} \tag{14.19}$$

where λ is in units of nm and where n can take integer values begining with 3. For example, for n = 3, λ = 656.1 nm. This formula is now known as the *Balmer* formula and the series of lines that it fits is called the *Balmer series*. The wavelength 364.5 nm, corresponding to $n \to \infty$, is called the *series limit*. It was soon discovered that within experimental error, the wavelengths of the then known spectral lines in the hydrogen spectrum could be fit with a similar formula of the form

$$\lambda = a \frac{n_2^2}{n_2^2 - n_1^2} \tag{14.20}$$

where a = 364.5 nm, $n_1 = 2$ and n_2 = 3, 4, 5, 6….., for the first, second, third, ……… members of the series, the first four values of n_2 being those corresponding to H_α, H_β, H_γ, H_δ. By inverting both sides of eqn. (14.20), the Balmer's formula can be rewritten in slightly different form as

$$\frac{1}{\lambda} = \frac{1}{a} - \frac{n_1^2}{a n_2^2}$$

Putting $\frac{1}{\lambda} = \bar{\nu}_n$ the wave number (number of waves per m) and $\frac{n_1^2}{a}$ = R, this gives

$$\bar{\nu}_n = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{14.21}$$

where R is known as the *Rydberg constant*. Eqn. (14.21) represents what is known as the hydrogen series formula, which gives the wave number, $\bar{\nu}_n$, as the difference of two terms. The first term $\frac{R}{n_1^2}$ is the constant term (as n_1 has a constant value 2) and $\frac{R}{n_2^2}$ is the *variable term*. $\frac{R}{n_1^2}$ at the same time gives the series limit corresponding to $n_2 = \infty$.

The representation of $\overline{\nu}_n$ as the difference of two terms as given by eqn. (14.21) corresponds to the view of the wave number as the difference of level between two energy steps. The terms therefore represent the energy of the atom in its initial and final states.

14.4 Rydberg-Ritz combination principle

Another interesting property of the hydrogen wavelengths is summarized in the Rydberg-Ritz combination principle. It occurred to Rydberg that combinations of terms other than those giving the four chief series might correspond to spectral lines observed to be present in the spectra but not belonging to the series. This idea of Rydberg was generalised by Ritz in 1908 into a principle which achieved remarkable results in the classification of spectral lines. It may be stated thus :

"By a combination of the terms that occur in the Rydberg or Balmer formula, other relations can be obtained holding good for new lines and even new series".

For instance, series other than that of Balmer in the hydrogen spectrum were predicted even before they were actually discovered. The first two lines of the Balmer series, H_α and H_β are represented by

$$\overline{\nu}_\alpha = R\left[\frac{1}{2^2}-\frac{1}{3^2}\right] \text{ and } \overline{\nu}_\beta = R\left[\frac{1}{2^2}-\frac{1}{4^2}\right]$$

Combining, we get

$$\overline{\nu}_\beta - \overline{\nu}_\alpha = R\left[\left(\frac{1}{2^2}-\frac{1}{4^2}\right)-\left(\frac{1}{2^2}-\frac{1}{3^2}\right)\right]$$

$$= R\left(\frac{1}{3^2}-\frac{1}{4^2}\right)$$

This represents a new line – indeed, the first line of a new series in the infra-red, discovered by Paschen and consequently known as *Paschen series.* Similarly the second line of the same series can be obtained by forming the difference of H_γ and H_α and so on. In the same manner, another series, also in the infra-red, discovered by Brackett, can be obtained with the combination principle.

Giving n_1 the values of 1 to 5 in eqn. (14.21) the following formulae, each representing a different series of the hydrogen spectrum, are obtained :

$$\text{Lyman series : } \bar{v}_n = R\left[\frac{1}{1^2}-\frac{1}{n_2^2}\right] \quad n_2 = 2, 3, 4, \ldots\ldots \tag{14.22}$$

$$\text{Balmer series : } \bar{v}_n = R\left[\frac{1}{2^2}-\frac{1}{n_2^2}\right] \quad n_2 = 3, 4, 5, \ldots\ldots \tag{14.23}$$

$$\text{Paschen series : } \bar{v}_n = R\left[\frac{1}{3^2}-\frac{1}{n_2^2}\right] \quad n_2 = 4, 5, 6, \ldots\ldots \tag{14.24}$$

$$\text{Brackett series : } \bar{v}_n = R\left[\frac{1}{4^2}-\frac{1}{n_2^2}\right] \quad n_2 = 5, 6, 7, \ldots\ldots \tag{14.25}$$

$$\text{Pfund series : } \bar{v}_n = R\left[\frac{1}{5^2}-\frac{1}{n_2^2}\right] \quad n_2 = 6, 7, 8, \ldots\ldots \tag{14.26}$$

The first series, as given by eqn. (14.22), was discovered by Lyman in the extreme ultra-violet region. The third series, given by eqn. (14.24), was discovered by Paschen as mentioned above. The fourth and fifth, given by eqns. (14.25) and (14.26) were discovered when they were predicted by Brackett and Pfund respectively.

***Example 14.3** The series limit of the Paschen series ($n_1 = 3$) is 820.1 nm. What are the three longest wavelengths of the Paschen series?*

Soln.

From eqn. (14.20)

$$\lambda = a\frac{n_2^2}{n_2^2 - n_1^2}$$

we have, for Paschen series, $a = 820.1$ nm, $n_1 = 3$.

For the three longest wavelengths $n_2 = 4, 5$ and 6.

14.5 The Bohr Atom Model

In 1913, Bohr suggested a model of the atom for which he was awarded Nobel Prize for Physics in 1922. He assumed basically Rutherford nuclear-model of the atom and tried to overcome the defects. His theory is a combination of the ideas of classical physics and the principle of quantum theory of radiation as enunciated by Planck and Einstein. Following Rutherford's proposal that the mass and positive charge of the atom are concentrated in a very small region at the centre of the atom, Bohr suggested that the atom was in fact like a miniature planetary system, with the electrons circulating about the nucleus like planets circulating about the sun. The atom thus does not collapse under the influence of the electrostatic coulomb attraction between nucleus and electrons for the same reason that the solar system doesn't collapse under the influence of the gravitational attraction between sun and planets. In both cases, the attractive force provides the centripetal acceleration necessary to maintain the orbital motion. Bohr atom model is based on the following postulates :

First postulate : The first postulate, referring to the electronic structure of the atom, states that the atom consists of a positively charged nucleus at the centre. The negatively charged particles known as the electrons move round the nucleus in various orbits just as the planets move round the sun. The electrons, however, cannot revolve in all possible orbits as suggested by the classical theory, but only in certain definite orbits satisfying quantum conditions. The motion of the electron in these orbits, though governed by the ordinary laws of mechanics and electrostatics, is not subject to the electromagnetic theory demanding continuous radiation of energy.

These orbits may therefore be considered as *privileged orbits*; they may also be treated as *stable* and *stationary*. These privileged quantum orbits are the *non-radiating* paths of the electrons and are also referred to as energy levels. Thus the difficulty with regard to the stability of the atom is overcome.

Second postulate : The second postulate concerning the *origin of spectral lines* states that the radiation of energy takes place only when an electron jumps from one permitted orbit to another of lower energy. The energy thus radiated, which is equal to the difference in the energies of the two orbits involved must be a quantum of energy hν.

Bohr took, for the application of these ideas, hydrogen, the simplest of all elements, which had already been investigated extensively and was known to contain only one electron. He tackled the problems of electronic structure and the origin of the spectral line connected with the hydrogen atom as follows :

Electronic Structure

Nature of the privileged quantum orbits

Let us consider the general case of a linear simple harmonic oscillator. Its displacement x at any instant of time t is given by

$$x = A \sin 2\pi\nu t \tag{14.27}$$

where A is the amplitude and ν the frequency. As the total energy of the oscillator changes from all kinetic at the equilibrium position to all potential at the maximum displacement, it can be determined by computing the kinetic energy at the equilibrium position.

The kinetic energy of the oscillator at any instant t = $\frac{1}{2}$m $(dx/dt)^2$ where m is the mass of the oscillator and dx/dt its linear velocity at the instant considered.

Now

$$\frac{dx}{dt} = 2\pi\nu A \cos 2\pi\nu t$$

At the equilibrium position dx/dt is maximum.

Hence $\left(\frac{dx}{dt}\right)_{max} = 2\pi\nu A$

since the maximum value of cos $2\pi\nu tA$ is 1.

∴ The total energy = maximum kinetic energy

$$= \frac{1}{2} m \left(\frac{dx}{dt}\right)^2_{max}$$

$$= \frac{1}{2} m\, (2\pi\nu A)^2$$

$$= 2\pi^2\nu^2A^2m.$$

According to the quantum theory, this energy should be an integral multiple of $h\nu$.

$$\therefore \quad nh\nu = 2\pi^2\nu^2A^2m$$

$$\text{or,} \quad nh = 2\pi^2A^2\nu m \qquad (14.28)$$

The momentum p_x of the oscillator at the instant t is given by

$$p_x = m\left(\frac{dx}{dt}\right) = m2\nu\pi A \cos 2\pi\nu t$$

Putting $m.2\pi\nu A = B$, $p_x = B \cos 2\pi\nu t$

$$\therefore \quad \frac{p_x}{B} = \cos 2\pi\nu t$$

From eqn. (14.27) $\frac{x}{A} = \sin 2\pi\nu t$

$$\therefore \quad \frac{x^2}{A^2} + \frac{p_x^2}{B^2} = 1$$

The above expression means that the relation between p_x and x is given by an ellipse. If a graph is drawn with x as abscissa and p_x as ordinate, it will be an ellipse whose semi-major and semi-minor axes are A and B respectively. Considering such an elipse (Fig. 14.11) let dx be the width of an element at a distance x from the origin. Let p_x be the value of the ordinate corresponding to x. Then the area of the

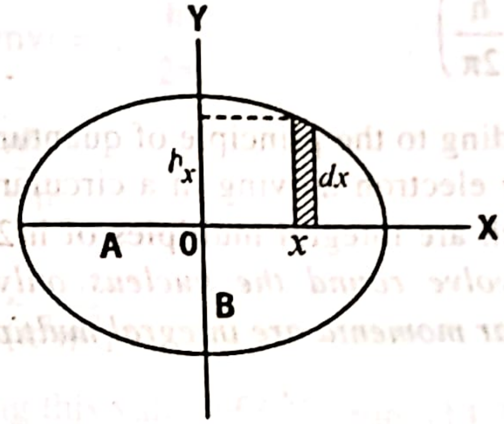


Fig. 14.11

element considered is $p_x.dx$. The area of the ellipse is obtained by the integration of $p_x.dx$ over a complete cycle.

$$\text{or,} \quad \oint p_x.dx = \text{area of the ellipse}$$

$$= \pi \times A \times B$$

$$= 2\pi^2 A^2 \nu m$$

$$= nh \qquad \text{from eqn. (14.28)}$$

The integration $\oint p_x.dx$ is known as the phase integral. Thus the phase integral of a linear oscillator is an integral multiple of h, the Planck's constant. Now an electron in uniform circular motion is equivalent to a harmonic oscillator. Thus the above relation can also be applied to an orbiting electron. Replacing the linear momentum p_x by the angular momentum p_φ and the linear element dx by the angular element $d\varphi$, we can write

$$\oint p_\varphi.d\varphi = nh \qquad (14.29)$$

But $p_\varphi = Iw$, where I is the moment of inertia of the electron and w the angular velocity. The angular velocity being assumed constant, p_φ is also constant.

$$\therefore \quad \oint p_\varphi.d\varphi = p_\varphi.\int_0^{2\pi} d\varphi = p_\varphi.2\pi = nh$$

$$\text{or,}\quad p_\varphi = n\left(\frac{h}{2\pi}\right)$$

Hence according to the principle of quantum theory, the angular momentum of the electron moving in a circular orbit can have only those values which are integral multiples of $h/2\pi$, or conversely *the electrons can revolve round the nucleus only in those orbits in which their angular momenta are integral multiples of* $h/2\pi$.

Radii of the permitted orbits

Consider an atom whose nucleus has a positive charge Ze and mass M. Let an electron of charge (– e) and mass m move round the nucleus in an orbit of radius r. Since $M >> m$, the nucleus is stationary. Hence the mass of the nucleus does not come into calculation.

The electrostatic force of attraction between the nucleus and the electron

$$= \frac{1}{4\pi\epsilon_0}\frac{Ze.e}{r^2} = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r^2}$$

The centrifugal force of repulsion between the two resulting from the circular motion of the electron

$$= \frac{mv^2}{r}$$

The system will be stable if

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r^2} \qquad (14.29)$$

Now the angular momentum of the electron moving in a circular orbit of radius r,

$$Iw = mr^2.w = mr^2.\frac{v}{r} = mvr$$

According to Bohr's first postulate,

$$Iw = mvr = n.\frac{h}{2\pi}$$

$$\text{or,}\quad v = \frac{nh}{2\pi mr} \tag{14.30}$$

$$\text{or,}\quad v^2 = \frac{n^2h^2}{4\pi^2m^2r^2}$$

Substituting this value of v^2 in eqn. (14.29)

$$\frac{m}{r}\left(\frac{n^2h^2}{4\pi^2m^2r^2}\right) = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r^2}$$

$$\text{or,}\quad r = \frac{n^2h^2\epsilon_0}{\pi Ze^2m} \tag{14.31}$$

From eqn. (14.31) we see that $r \propto n^2$. The radii of the orbits are in the ratio of 1 : 4 : 9 : 16 : 25 etc. The radius of the second orbit is four times the radius of the first orbit.

For hydrogen, $Z = 1$

$\therefore$ Radius of the n^{th} permitted orbit for hydrogen,

$$r_n = \frac{n^2h^2\epsilon_0}{\pi e^2m} \tag{14.32}$$

The radius of the first orbit ($n = 1$) for the hydrogen atom

$$r_1 = \frac{1^2\times(6.625\times10^{-34})^2(8.854\times10^{-12})}{\pi(1.6\times10^{-19})^2(9.11\times10^{-31})}$$

$$= 5.29 \times 10^{-11}\ \text{m}$$

$$= 0.529\ \text{A.U.}$$

The radius of the first orbit of the hydrogen atom is called the *Bohr radius*. The diameter of the first orbit is, therefore, of the order of 10^{-10}m, which agrees with the values of the diameters of atoms computed by various other methods.

Orbital velocity of electron

Substituting the value of r as given by eqn. (14.31) in eqn. (14.29), we get

$$v = \frac{Ze^2}{2\epsilon_0 nh} \qquad (14.33)$$

Eqn. (14.33) shows that the velocity of the electron decreases with the increase in the order of the orbit. The velocity of the electron in the first orbit of the hydrogen atom is given by

$$v = \frac{(1.6\times10^{-19})^2}{2\times8.854\times10^{-12}\times6.625\times10^{-34}} \qquad (Z = 1)$$

$$= 2.2 \times 10^6 \text{ m/s.}$$

Orbital energy of the electron

The total energy of the electron in any orbit is the sum of its kinetic and potential energies. The potential energy of the electron is considered to be zero when it is at an infinite distance from the nucleus. Potential energy of an electron in an orbit is given by the work done in bringing the electron from infinity to that orbit. The amount of work done is obtained by integrating the electron static force of attraction between the nucleus and the electron from the limits ∝ to r.

Potential energy (P.E.) of the electron

$$= \int_{\infty}^{r} \frac{Ze^2}{4\pi\epsilon_0} \frac{dr}{r^2} = \frac{-Ze^2}{4\pi\epsilon_0 r} \qquad (14.34)$$

Kinetic energy (K.E.) of the electron

$$= \frac{1}{2}mv^2 = \frac{Ze^2}{8\pi\epsilon_0 r} \quad \text{(from eqn. 14.29)} \qquad (14.35)$$

The potential energy is negative because energy must be given to the electron to bring it far away from the nucleus to zero energy level.

∴ The total energy of the electron in the nth orbit

$$= E_n = \text{P. E.} + \text{K. E.}$$

$$= -\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{Ze^2}{8\pi\epsilon_0 r} = -\frac{Ze^2}{8\pi\epsilon_0 r}$$

Substituting the value of r from eqn. (32)

$$E_n = -\frac{me^4Z^2}{8\epsilon_0^2 n^2h^2} \tag{14.36}$$

where n = 1, 2, 3, ... etc., for the energy states that it is possible for the electron to have. Negative values of energy imply that the electron is bound to the nucleus by attractive forces so that energy must be supplied to the electron in order to separate it completely from the nucleus.

The integer n is called the *total* or *principal* quantum number and it can have any of the series of values 1, 2, 3, ... As can be seen from eqn. (14.36), $E_n \propto 1/n^2$ *i.e.*, the values of n determine the energies of the state. When n is large, the energy is large *i.e., less negative* than for small integers. The energy required to remove an electron from a particular state to infinity is called the *binding energy* of that state. It is numerically equal to E_n and is also sometimes referred to as *work function.*

Substituting the numerical values of the symbols of eqn. (14.36) we get

$$E_n = -\frac{(9.11\times10^{-31})(1.6\times10^{-19})^4}{(8)\,(8.85\times10^{-12})^2 n^2 (6.62\times10^{-34})^2} \text{ Joule}$$

(Z = 1 for hydrogen atom)

Since 1eV = 1.6 x 10^{-19} J,

$$E_n = -\frac{(9.11\times10^{-31})\,(1.6\times10^{-19})^4}{(8)\,(8.854\times10^{-12})^2\,(6.625\times10^{-34})^2 n^2\,(1.6\times10^{-19})} \text{ eV}$$

$$= -\frac{13.6}{n^2} \text{ eV.}$$

Thus, in the case of hydrogen atom the binding energy of the electron in the n^{th} orbit is given by $\frac{13.6}{n^2}$ eV. Energy required to extricate the electron from inner orbits is therefore greater than from outer orbits.

Origin of spectral lines

According to Bohr's second postulate, even though the electron does not radiate when it remains in any particular stationary state (permitted orbit), it can emit radiation when it moves to a lower energy level. Suppose an electron jumps from an outer initial orbit n_2 of higher energy to an inner orbit n_1 of lower energy, then the energy difference between the two levels appears as a quantum of radiation whose energy hν is equal to the energy difference between the two levels. Or

$$h\nu = E_{n_2} - E_{n_1}$$

So, the frequency of emitted radiation

$$\nu = \frac{E_{n_2} - E_{n_1}}{h}$$

So, for hydrogen atom (Z = 1),

$$E_{n_2} = -\frac{me^4}{8\epsilon_0^2 h^2} \cdot \frac{1}{n_2^2}$$

$$E_{n_1} = -\frac{me^4}{8\epsilon_0^2 h^2} \cdot \frac{1}{n_1^2}$$

$$\therefore \quad \nu = \frac{E_{n_2} - E_{n_1}}{h} = \frac{me^4}{8\epsilon_0 h^3}\left\{\frac{1}{n_1^2} - \frac{1}{n_2^2}\right\} \tag{14.38}$$

The wavelength of the emitted radiation is

$$\lambda = \frac{c}{\nu} = \frac{8\epsilon_o h^3 c}{me^4}\left(\frac{n_1^2 n_2^2}{n_2^2 - n_1^2}\right)$$

The wave number $\bar{\nu}$ of a radiation is defined as the reciprocal of its wavelength λ in vacuum and gives the number of waves contained in unit length in vacuum.

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\nu}{c}$$

$$= \frac{me^4}{8\epsilon_0^2 ch^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{14.39}$$

The numerical coefficient in eqn. (14.39) is called the *Rydberg constant* R_∞

or, $$R_\infty = \frac{me^4}{8\epsilon_0^2 ch^3}$$

Substituting the values of m, e, ϵ_0, c and h

$$R_\infty = \frac{(9.11\times10^{-31})(1.6\times10^{-19})^4}{(8)(8.854\times10^{-12})^2(3\times10^8)(6.625\times10^{-34})^3}$$

$$= 1.097 \times 10^7 \text{ m}^{-1}$$

$$\therefore \quad \bar{\nu} = R_\infty\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

Eqn. (14.39) suggests a very simple picture of the origin of spectral lines. A spectral line is emitted when an electron initially revolving in an orbit of quantum number n_2 drops to an inner orbit of quantum number n_1, the wave number being given by the above relation. A whole series of lines corresponds to a given inner orbit. The different series observed in the hydrogen spectrum are as under:

(1) Lyman Series

It consists of all those wavelengths which are emitted when an electron jumps from different higher orbits to the first orbit (K-orbit). Here $n_1 = 1$ and $n_2 = 2, 3, 4 \ldots\ldots\ldots\ldots \infty$

The wavelengths of different members of this series may be found from the following relation:

$$\bar{\nu} = \frac{1}{\lambda} = R\left(\frac{1}{1^2} - \frac{1}{n_2^2}\right) = R\left(1 - \frac{1}{n_2^2}\right)$$

(i) First member : $n_2 = 2$

$$\frac{1}{\lambda} = R\left(1 - \frac{1}{2^2}\right) = \frac{3R}{4}$$

or, $$\lambda = \frac{4}{3R} = \frac{4}{3\times1.097\times10^7}\text{ m} = 1216 \times 10^{-10}\text{m} = 1216 \text{ Å}.$$

(ii) Second member : $n_2 = 3$

$$\frac{1}{\lambda} = R\left(1 - \frac{1}{3^2}\right) = \frac{8R}{9}$$

$$\therefore \quad \lambda = \frac{9}{8R} = \frac{9}{8 \times 1.097 \times 10^7} \text{ m} = 1026 \text{ Å}$$

And so on.

(iii) Limiting number : $n_2 = \infty$

The limiting member of the series *i.e.*, the shortest wavelength of the series is obtained by putting $n_2 = \infty$ in the above relation.

$$\therefore \quad \frac{1}{\lambda} = R\left(1 - \frac{1}{\infty^2}\right) = R$$

$$\therefore \quad \lambda = \frac{1}{R} = \frac{1}{1.097 \times 10^7} \text{ m} = 912 \text{ Å} = 91.2 \text{ nm}$$

This series of wavelengths lies in the ultraviolet region of the spectrum. The jumps corresponding to this series are shown in Fig.14.12.

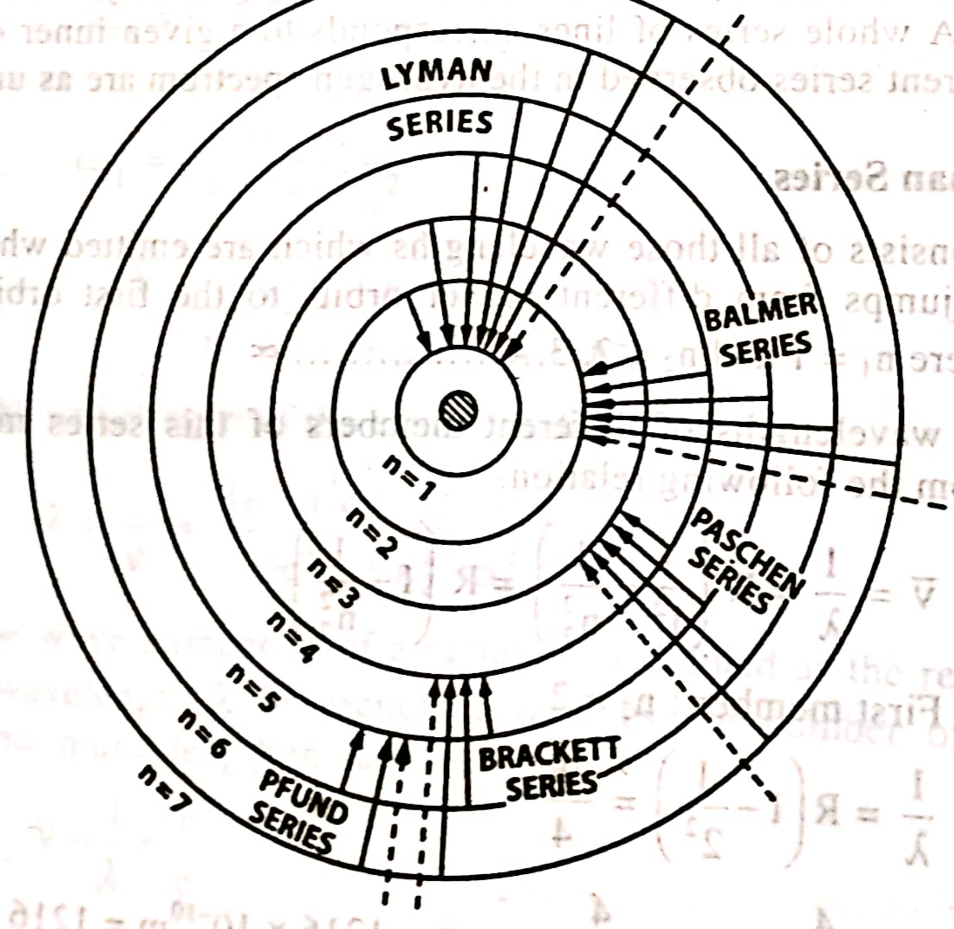


Fig. 14.12

(2) Balmer Series

The series originates when an electron jumps from different higher orbits to the second orbit (L – orbit). Here $n_1 = 2$ and $n_2 = 3, 4, 5, \ldots\ldots\ldots \infty$.

The wavelengths of the different members of the series may be obtained from the following relation

$$\bar{\nu} = \frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n_2^2}\right) = R\left(\frac{1}{4} - \frac{1}{n_2^2}\right)$$

The wavelengths of the different members are as under:

(i) First member : $n_2 = 3$

$$\therefore \quad \frac{1}{\lambda} = R\left(\frac{1}{4} - \frac{1}{3^2}\right) = \frac{5R}{36}$$

$$\therefore \quad \lambda = \frac{36}{5R} = \frac{36}{5 \times 1.097 \times 10^7} \text{ m} = 6563 \times 10^{-10} \text{ m} = 6563 \text{ Å}.$$

Similarly the wavelength of the second member of the series is obtained by putting $n_2 = 4$

or, $\lambda = 4861$ Å

and so on.

The wavelength of the limiting member of the series is obtained by putting $n_2 = \infty$

or, $\lambda = 3646$ Å $= 364.6$ nm.

This series lies in the visible and under ultraviolet region of the solar spectrum.

(3) Paschen Series

The different wavelengths of this series are given by

$$\frac{1}{\lambda} = R\left(\frac{1}{9} - \frac{1}{n_2^2}\right) \text{ where } n_2 = 4, 5, 6, \ldots\ldots\ldots \propto$$

The series lies in the infra-red region.

(4) Brackett Series

If $n_1 = 4$ and $n_2 = 5, 6, 7, \ldots\ldots\ldots \propto$, we get the Brackett series.

$$\frac{1}{\lambda} = R\left(\frac{1}{4^2} - \frac{1}{n_2^2}\right) \text{ where } n_2 = 5, 6, 7, \ldots\ldots\ldots \propto$$

(5) Pfund Series

If $n_1 = 5$ and $n_2 = 6, 7, 8, \ldots\ldots\ldots \propto$, we get Pfund series.

$$\frac{1}{\lambda} = R\left(\frac{1}{5^2} - \frac{1}{n_2^2}\right) \text{ where } n_2 = 6, 7, 8, \ldots\ldots\ldots \propto$$

Brackett and Pfund series lie in the very far infra red region of the hydrogen spectrum. By putting $n_2 = \propto$ in each of the series, the wavelength of the limiting member of the series can be obtained.

The energy-level diagram

The discovery of sharp energy states of the hydrogen atom was a finding that had very far-reaching consequences in physics. All atomic, molecular and nuclear systems have discrete energy levels. These levels can be represented graphically as shown in Fig. 14.13. The quantum numbers are shown at the left and the corresponding energies of hydrogen in electron volts are given at the right. In this array of energies, the higher (less negative) energies are at the top, while the

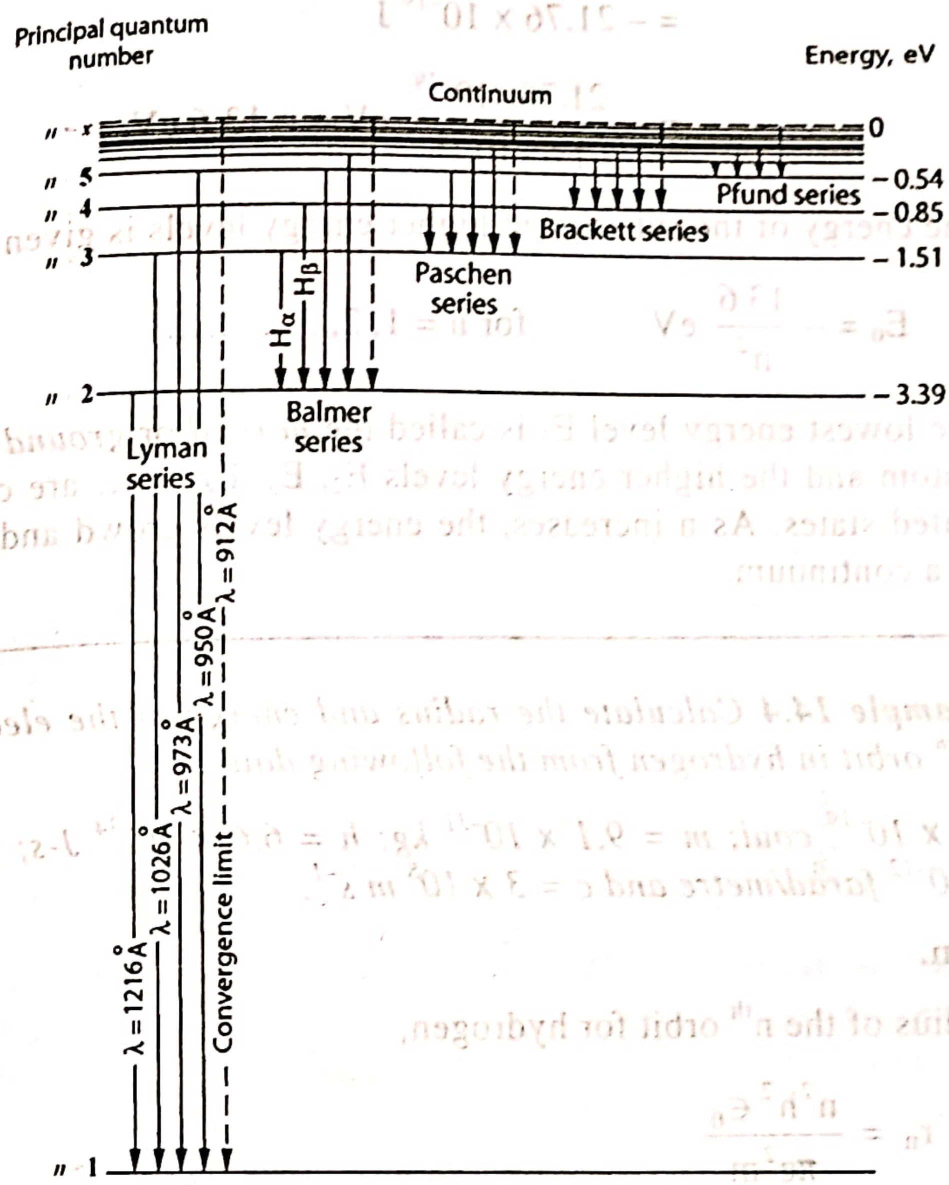


Fig. 14.13

lower (more negative) toward the bottom. In a normal unexcited hydrogen atom, the electron is in its lowest energy state at the bottom, with n = 1. In the expression

$$E_n = -\frac{me^4Z^2}{8\epsilon_0^2 n^2h^2}$$

putting Z = 1 and n = 1 for hydrogen atom and substituting the values of m, e, ϵ_0 and h, the energy of this lowest energy state is

$$-\frac{me^4Z^2}{8\epsilon_0^2 n^2h^2} = -\frac{(9.11\times10^{-31})(1.6\times10^{-19})^4\,(1)^2}{8(8.854\times10^{-12})^2\,(1)^2\,(6.625\times10^{-34})^2}$$

$$= -21.76 \times 10^{-19} \text{ J}$$

$$= -\frac{21.76\times10^{-19}}{1.6\times10^{-19}} \text{ eV} = -13.6 \text{ eV.}$$

The energy of the subsequent higher energy levels is given by

$$E_n = -\frac{13.6}{n^2} \text{ eV} \qquad \text{for } n = 1, 2, 3 \ldots\ldots\ldots$$

The lowest energy level E_1 is called the *normal* or *ground state* of the atom and the higher energy levels E_2, E_3, E_4, ... are called the excited states. As n increases, the energy levels crowd and tend to form a continuum.

***Example 14.4** Calculate the radius and energy of the electron in the n^{th} orbit in hydrogen from the following data :*

$e = 1.6 \times 10^{-19}$ coul; $m = 9.1 \times 10^{-31}$ kg; $h = 6.6 \times 10^{-34}$ J-s; $\epsilon_0 = 8.85 \times 10^{-12}$ farad/metre and $c = 3 \times 10^8$ m s^{-1}.

Soln.

Radius of the n^{th} orbit for hydrogen,

$$r_n = \frac{n^2 h^2 \epsilon_0}{\pi e^2 m}$$

$$= \frac{n^2 (6.6\times10^{-34})^2 (8.85\times10^{-12})}{(3.14)(1.6\times10^{-19})^2 (9.1\times10^{-31})}$$

$$= 0.527 \times 10^{-10}\, n^2 \text{ metre}$$

$$= 0.527\, n^2 \text{ A.U.}$$

Total energy of the electron in the n^{th} orbit,

$$E_n = \frac{-me^4}{8\epsilon_0^2 n^2 h^2}$$

$$= -\frac{(9.1\times10^{-31})(1.6\times10^{-19})^4}{8(8.85\times10^{-12})^2 n^2 (6.6\times10^{-34})^2}$$

$$= -\frac{21.85\times10^{-19}}{n^2}\ J$$

$$= -\frac{21.85\times10^{-19}}{n^2(1.6\times10^{-19})}\ eV$$

$$= -\frac{13.66}{n^2}\ eV.$$

Example 14.5 (a) What is the wavelength of the least energetic photon in the Balmer spectrum? (b) What is the wavelength of the series limit for the Balmer series.

Soln.

(a) The relation for Balmer series is given by

$$\bar{\nu} = \frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

where $n_1 = 2$ and $n_2 = 3, 4, 5, \ldots\ldots\ldots$

From the relation E = hν, the least energetic photon has the smallest frequency and thus the greatest wavelength. This means that n_2 must be equal to 3 in the above expression; any higher value of n_2 will yield a smaller wavelength. With these substitutions, we have

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{3^2}\right)$$

$$= (1.097 \times 10^7\ m^{-1})\left(\frac{1}{2^2} - \frac{1}{3^2}\right)$$

$$= 1.524 \times 10^6\ m^{-1}$$

$\therefore\quad \lambda = 6.563 \times 10^{-7}\ m = 656.3\ nm.$

(b) To find the series limit, we put $n_2 = \infty$.

$$\therefore\quad \frac{1}{\lambda} = (1.097\times10^7\ m^{-1})\left(\frac{1}{2^2} - \frac{1}{\infty^2}\right)$$

$$= \left(1.097\times10^{7}\,m^{-1}\right)\left(\frac{1}{2^2}-0\right)$$

$$= 2.743 \times 10^{6}\ m^{-1}$$

$$\therefore\quad \lambda = 3.646 \times 10^{-7} m = 364.6 m.$$

***Example 14.6** Show that the velocity of the electron in the first Bohr orbit is (1/137)c where c is the velocity of light.*

Soln.

The expression for the velocity of electron in a given orbit is given by (eqn. 14.33)

$$v = \frac{Ze^2}{2\epsilon_0\, nh} \qquad \text{Here } Z = 1$$

$$\therefore\quad v = \frac{\left(1.6\times10^{-19}\right)^2}{(2)\left(8.854\times10^{-12}\right)\left(6.6\times10^{-34}\right)}$$

$$= 2.19 \times 10^{6}\ m/s$$

$$\therefore\quad \frac{v}{c} = \frac{2.19\times10^{6}}{3\times10^{8}} = \frac{1}{137}$$

***Example 14.7** Determine the separation of the first line of the Balmer series in a spectrum of ordinary hydrogen and tritium (mass number 3). Given that $R_\infty = 10973940\ m^{-1}$.*

Soln.

Rydberg constant for ordinary hydrogen atom is

$$R_H = \frac{R_\infty}{(1+m/M)}$$

Rydberg constant for tritium is

$$R_T = \frac{R_\infty}{(1+m/3M)}$$

Let λ_1 and λ_2 be the wavelengths of the first line of Balmer series of hydrogen and tritium respectively.

$$\lambda_1 = R_H\left(\frac{1}{2^2}-\frac{1}{3^2}\right)=\frac{5}{36}R_H; \quad \text{or} \quad \lambda_1 = \frac{36}{5R_H}$$

$$\frac{1}{\lambda_2} = R_T\left(\frac{1}{2^2}-\frac{1}{3^2}\right)=\frac{5}{36}R_T; \quad \text{or} \quad \lambda_2 = \frac{36}{5R_T}$$

$$\therefore \quad \lambda_1 - \lambda_2 = \frac{36}{5}\left(\frac{1}{R_H}-\frac{1}{R_T}\right)$$

$$= \frac{36}{5}\left[\frac{(1+m/M)}{R_\infty}-\frac{(1+m/3M)}{R_\infty}\right]$$

$$= \frac{36}{5R_\infty}[(1+m/M)-(1+m/3M)]$$

$$= \frac{36}{5R_\infty}\left[\frac{2m}{3M}\right]$$

$$= \frac{36}{5\times10973740}\times\frac{2}{3}\times\frac{1}{1836}$$

$$= 2.38\times10^{-10}\ \text{m}$$

$$= 2.38\ \mathring{A}.$$

***Example 14.8** The series limit of the Paschen series (n_l = 3) is 820.1 nm. What are the three longest wavelengths of the Paschen series ?*

Soln.

From eqn. 14.20

$$\lambda = (820.1)\frac{n_2^2}{n_2^2-n_1^2} \qquad (n_2 = 4, 5, 6, \ldots\ldots)$$

$$N_2 = 4 \qquad \lambda = (820.1)\frac{4^2}{4^2-3^2} = 1875\ \text{nm}$$

$$n_2 = 5 \qquad \lambda = (820.1)\frac{4^2}{5^2 - 3^2} = 1281 \text{ nm}$$

$$n_2 = 6 \qquad \lambda = (820.1)\frac{4^2}{6^2 - 3^2} = 1094 \text{ nm}$$

***Example 14.9** In a Frank-Hertz experiment, hydrogen was bombarded by 10.21 eV and 12.10 eV electrons which resulted in the emission of three spectral lines. Calculate their wavelengths.*

Soln.

The energy states of hydrogen atom are – 13.6 eV (n = 1), – 3.4 eV (n = 2), – 1.51 eV (n = 3), – 0.85 eV (n = 4), etc. The three spectral lines obviously correspond to transition from n = 2 to n = 1 (13.6 – 3.4 = 10.2 eV), n = 3 to n = 1 (13.6 – 1.51 = 12.10 eV) and n = 3 to n = 2 (3.4 – 1.51 = 12.10 – 10.21 = 1.89 eV).

(i) $n = 2 \rightarrow n = 1$

The wave number $\overline{\nu}$ is given by

$$\overline{\nu} = \frac{1}{\lambda} = R_H\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

where R_H is the Rydberg's constant for hydrogen

($R_H = 1.09678 \times 10^7 \text{ m}^{-1}$)

$$\therefore \quad \overline{\nu} = \frac{1}{\lambda} = 1.09678 \times 10^7\left(\frac{1}{1^2} - \frac{1}{2^2}\right)$$

$$= (1.09678 \times 10^7)(0.75)$$

$$= 0.822585 \times 10^7.$$

$$\therefore \quad \lambda = \frac{1}{0.822585 \times 10^7} = 1.21568 \times 10^{-7} \text{ m}$$

$$= 1215 \times 10^{-10} = 1215 \text{ Å}$$

(ii) $n = 3 \rightarrow n = 2$

$$\overline{\nu} = \frac{1}{\lambda} = 1.09678 \times 10^7\left(\frac{1}{2^2} - \frac{1}{3^2}\right)$$

$$= 0.9749 \times 10^7 \text{ m}^{-1}.$$

$$\therefore \quad \lambda = \frac{1}{0.9749 \times 10^7} \text{ m} = 1.025 \times 10^{-7}\text{m}.$$

$$= 1025 \times 10^{-10} = 1025 \text{ Å}.$$

(iii) $n = 3 \rightarrow n = 1$

$$\bar{\nu} = \frac{1}{\lambda} = 1.09678 \times 10^7 \left(\frac{1}{1^2} - \frac{1}{3^2}\right)$$

$$= 0.1523 \times 10^7 \text{ m}^{-1}$$

$$\therefore \quad \lambda = \frac{1}{0.1523 \times 10^7} \text{ m} = 6.566 \times 10^{-7}\text{m}.$$

$$= 6566 \times 10^{-10} = 6566 \text{ Å}.$$

***Example 14.10** At what speed must the electron revolve round the nucleus of a hydrogen atom in order that it may not be pulled into the nucleus by electrostatic attraction? Calculate the time taken by the electron to complete one revolution. radius of the orbit = 0.5×10^{-10} m, $m = 9.1 \times 10^{-31}$ kg, $e = 1.6 \times 10^{-19}$ coulomb.*

Soln.

The electrostatic force of attraction between the nucleus and the electron $= \dfrac{1}{4\pi\epsilon_0} \dfrac{(Ze)(-e)}{r^2}$

r is the radius of the orbit in which the electron moves.

The centrifugal force on the electron $= \dfrac{mv^2}{r}$

The system will be stable if

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r^2}$$

$$\text{or,} \quad v^2 = \frac{r}{4\pi\epsilon_0 m} \frac{Ze^2}{r^2} = \frac{Ze^2}{4\pi\epsilon_0 mr}$$

$$= \frac{(1)(1.6\times10^{-19})^2}{(4)(3.14)(8.85\times10^{-12})(9.1\times10^{-31})(0.5\times10^{-10})}$$

$$= \frac{2.56\times10^{-38}}{505.92\times10^{-53}}$$

$$= 5.06 \times 10^{-3} \times 10^{-38} \times 10^{53}$$

$$= 5.06 \times 10^{12}.$$

$\therefore\quad v = 2.249 \times 10^{6}\ m\ s^{-1}.$

The circumference of the orbit

$$= 2\pi r = (2)(3.141)(0.5 \times 10^{-10})$$

$$= 3.141 \times 10^{-10}\ m.$$

$\therefore$ time taken to complete one revolution,

$$t = \frac{2\pi r}{v} = \frac{3.14\times10^{-10}}{2.249\times10^{6}}\ s$$

$$= 1.4 \times 10^{-16}\ s.$$

14.6 Effect of nuclear motion on atomic spectra

In Bohr theory, it was assumed that the mass M of the nucleus was so great compared with the mass m of the electron that the nucleus remained fixed at the centre of the circular orbits. This is rigorously true only if the mass of the nucleus is infinite. If the nuclear mass M is not infinite, both the nucleus and its orbital electron revolve around a common centre of mass with same angular velocity w. Let N and e represent the nucleus and electron of mass M and m respectively (Fig. 14.14). Both of them are rotating about their

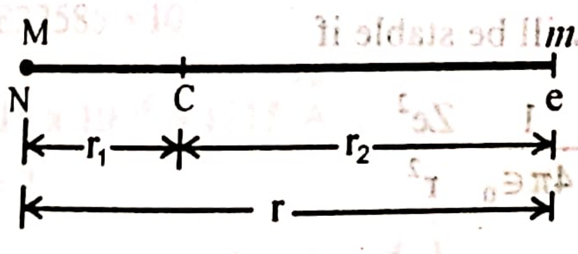


Fig. 14.14

common centre of mass C, the nucleus moving in a circle of radius r_1 and the electron in a circle of radius r_2. According to centre of mass theory

$$Mr_1 = mr_2 \quad (14.40)$$

Let r represent the distance between the nucleus and the electron. Then

$$r = r_1 + r_2 \quad (14.41)$$

$$\therefore \quad r_1 = r - r_2 = r - \frac{Mr_1}{m}$$

$$\text{or,} \quad r_1\left(1+\frac{M}{m}\right) = r$$

$$\text{or,} \quad r_1 = \left(\frac{m}{M+m}\right).r \quad (14.42)$$

$$\text{Similarly} \quad r_2 = \left(\frac{M}{M+m}\right).r \quad (14.43)$$

Since the two masses are revolving in circular orbits, there should be two separate equations as conditions of stability, *viz.*,

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = \frac{mv^2}{r_1} \quad \text{for the electron} \quad (14.44)$$

$$\text{and} \quad \frac{Ze^2}{4\pi\epsilon_0 r^2} = \frac{MV^2}{r_2} \quad \text{for the nucleus} \quad (14.45)$$

where the velocity of the electron, $v = wr_2$

and the velocity of the nucleus, $V = wr_1$

The total kinetic energy of the system is made up of two parts, that of the electron and that of the nucleus.

$\therefore$ Total K.E.= K.E. of the electron + K.E. of the nucleus

$$= \frac{1}{2}mv^2 + \frac{1}{2}MV^2$$

$$= \frac{1}{2} m w^2 r_2^2 + \frac{1}{2} M w^2 r_1^2$$

$$= \frac{1}{2} m w^2 \left(\frac{rM}{M+m} \right)^2 + \frac{1}{2} M w^2 \left(\frac{rm}{M+m} \right)^2$$

[from eqns. 14.42 and 14.43]

$$= \frac{1}{2} r^2 w^2 \frac{Mm}{(M+m)^2} (M+m)$$

$$= \frac{1}{2} \frac{Mm}{(M+m)} w^2 r^2$$

Putting $\frac{Mm}{(M+m)} = \frac{m}{1+m/M} = \mu$, we get

Total kinetic energy $= \frac{1}{2} \mu w^2 r^2$

where μ is called the *reduced mass* of the electron. μ tends to m as M tends to infinity.

If the nucleus was stationary, the total kinetic energy $= \frac{1}{2} m v^2$ $= \frac{1}{2} m w^2 r^2$. The only effect of the finite mass of the nucleus is that the mass of the electron (m) is replaced by reduced mass (μ). Therefore to take into account the finite mass of the nucleus, the mass of the electron (m) in all the Bohr formulae should be replaced by the reduced mass (μ).

Equation for the wave-numbers of the spectral lines of the atom

Without following through the entire deduction, the final expression for the energy-levels of the hydrogen atom, in which the nuclear mass is not infinite, can be written as,

$$E_n = -\frac{\mu e^4 Z^2}{8 \epsilon_0^2 h^2} \left(\frac{1}{n^2} \right) \qquad (14.46)$$

This is the same as eqn. 14.36, except that μ replaces the electron mass m. Using the frequency condition, when an electron jumps from the Bohr orbit n_2 to the Bohr orbit n_1, the frequency of the radiation emitted is

$$\nu = \frac{\mu e^4 Z^2}{8\epsilon_0^2 h^3}\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right) \quad (14.47)$$

$$\text{and} \quad \bar{\nu} = \frac{\mu e^4 Z^2}{8\epsilon_0^2 ch^3}\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right) \quad (14.48)$$

The Rydberg constant for any element of atomic number Z is given by

$$R_Z = \frac{\mu e^4}{8\epsilon_0^2 ch^3} = \frac{me^4}{8\epsilon_0^2 ch^3} \times \frac{M_Z}{(M_Z + m)}$$

$$= R_\infty \times \frac{1}{1+m/M_Z}$$

where $R_\infty = \dfrac{me^4}{8\epsilon_0^2 ch^3}$ is the value of the Rydberg constant when the nucleus is at rest *i.e.*, $M_Z = \infty$. M_Z is, of course, the mass of the nucleus of the element with atomic number Z. The constant R_∞ can be computed by substituting the values of the quantities involved; but they are not known with spectroscopic precision. It is therefore preferably estimated from the spectroscopically observed value of the Rydberg constant for hydrogen R_H = 109,67770 m^{-1} and the known value of $m/M_H = 1/1840$ using the relation

$$R_H = R_\infty \left(\frac{1}{1+m/M_H}\right)$$

The value of R_∞ thus found is 10973740 m^{-1}. Eqn. (14.48) can then be written as

$$\bar{\nu} = Z^2 R_Z\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right) \quad (14.49)$$

14.7 The Correspondence Principle

As mentioned above, the calculations of transition wavelengths in atomic hydrogen on the basis of Bohr's atom model are in excellent agreement with the wavelengths observed in the emission and absorption spectra. However, in order to obtain this agreement, Bohr had to introduce two postulates that are radical departures from classical physics. In particular, an accelerated charged particle must radiate energy according to classical electromagnetic theory. But in Bohr's atom model, although an electron is accelerated as it moves in a circular orbit, it does not radiate energy unless it jumps to another orbit. This is a completely different case than the one we had in the study of special theory of relativity. It may be recalled that relativity gives one expression for the kinetic energy, $K = E - E_O$, whereas classical physics gives us another, $K = \frac{1}{2}mv^2$. However, it was shown that when $v << c$, $E - E_0$ reduces to $\frac{1}{2}mv^2$. Thus the two expressions are really not very different – one is merely a special case of the other. The dilemma associated with the accelerated electron is not simply a matter of atomic (quantum) physics being a special case of classical physics. Either the accelerated charge radiates, or it does not ! To overcome this serious dilemma Bohr, in 1932, proposed the *correspondence principle*, which states that

Quantum theory must agree with classical theory in the limit in which classical theory is known to agree with experiment.

or equivalently

Quantum theory must agree with classical theory in the limit of large quantum numbers.

Let us apply this principle to Bohr atom. According to classical physics, an electric charge moving in a circle radiates at a frequency equal to its frequency of rotation. If the electron rotates in an orbit of radius r with a speed v, then the period of revolution, T is $2\pi r/v$. Now the speed v is given by $v = \sqrt{2K/m}$, where K is the kinetic energy. Hence

$$T = \frac{2\pi r}{\sqrt{2K/m}}$$

From eqn. (35), $K = \frac{1}{2}mv^2 = \frac{e^2}{8\pi\epsilon_0 r}$ (Z = 1 for hydrogen)

$$\therefore \quad T = \frac{\sqrt{16\pi^3 \epsilon_0 mr^3}}{e} \tag{14.50}$$

The frequency ν is the inverse of the period.

$$\nu = \frac{1}{T} = \frac{e}{\sqrt{16\pi^3 \epsilon_0 mr^3}} \tag{14.51}$$

Using eqn. (31) for the allowed orbits, we get

$$\nu_n = \frac{me^4}{32\pi^3 \epsilon_0^2 \hbar^3} \cdot \frac{1}{n^3} \tag{14.52}$$

where $\hbar = \frac{h}{2\pi}$

A classical electron moving in an orbit of radius r_n would radiate at this frequency ν_n.

If the radius of the Bohr atom is made so large that it went from a quantum-sized object (10^{-10}m) to a laboratory-sized object (10^{-3}m), the atom should behave classically. Since the radius increases with increasing n like n^2, this classical behaviour should occur for n in the range $10^3 - 10^4$. Let us calculate the frequency of the radiation emitted by such an atom when the electron drops from the orbit n to the orbit n – 1. According to eqn. (14.37), the frequency is

$$\nu = \frac{me^4}{64\pi^3 \epsilon_0^3 \hbar} \left(\frac{1}{(n-1)^2} - \frac{1}{n^2}\right)$$

$$= \frac{me^4}{64\pi^3 \epsilon_0^3 \hbar^3} \frac{2n-1}{n^2(n-1)^2} \tag{14.53}$$

If n is very large, then we can approximate n – 1 by n and 2n – 1 by 2n, which gives

$$\nu \cong \frac{me^4}{64\pi^3 \epsilon_0^3 \hbar^3} \frac{2n}{n^4} = \frac{me^4}{32\pi^3 \epsilon_0^2 \hbar^3} \frac{1}{n^3} \qquad (14.54)$$

Eqn. (14.54) is identical with eqn. (14.52) for the "classical" frequency. The "classical" electron spirals slowly in toward the nucleus, radiating at the frequency given by eqn. (14.52), while the "Quantum" electron jumps from the orbit n to the orbit n – 1 and then to the orbit n – 2, and so forth, radiating at the frequency given by the identical eqn. (14.54). When the circular orbits are very large, this jumping from one circular orbit to the next smaller one looks very much like a spiral.

In the region of large n, where classical and quantum physics overlap, the classical and quantum expressions for the radiation frequencies are identical. This is an example of the application of Bohr's correspondence principle, the applications of which go far beyond the Bohr atom. The correspondence principle is important in understanding how we get from the domain in which the laws of classical physics are valid to the domain in which the laws of quantum physics are valid.

The difference in the frequencies of transition from one quantum state to the next lower state calculated on the basis of classical mechanics and quantum mechanics is given in the table below.

Table

The correspondence principle and the hydrogen atom.

Quantum Number n	Frequency in Hz of transition to the next lower state (classical mechanics)	Frequency in Hz of transition to the next lower state (quantum mechanics)	Difference (%)
2	8.22×10^{14}	24.7×10^{14}	67
5	5.26×10^{13}	7.40×10^{13}	29
10	6.58×10^{12}	7.72×10^{12}	15
50	5.26×10^{10}	5.43×10^{10}	3.1
100	6.580×10^{9}	6.680×10^{9}	1.5
1,000	6.5797×10^{6}	6.5896×10^{6}	0.15
10,000	6.5797×10^{3}	6.5807×10^{3}	0.015
25,000	4.2110×10^{2}	4.2113×10^{2}	0.007
100,000	6.5798	6.5799	0.0007

14.8 Excitation and Ionisation of atoms

According to Bohr's atomic theory, in order that the atom may radiate, the electron must abandon its normal orbit and move temporarily to an outer orbit of greater energy. When this happens, the atom is said to be in an *excited* state. It is obvious that the shifting of an electron to a higher orbit demands a supply of energy to the atom. This process of increasing the *internal energy* of the atom is called *excitation of the atom.* If the excitation is so intense that the electron is removed from a given orbit to an infinite distance from the nucleus ($n = \infty$), the atom is left with a net positive charge and is then said to be *ionised.* This extreme type of excitation is called *ionization of the atom.* Since the atom must absorb sufficient energy to raise itself from the normal to the excited or ionized state, it is obvious that the process of excitation and ionization is an *absorption phenomenon.* The state of excitation lasts only for a very short time, of the order of 10^{-8} sec, after which the atom comes back to its normal state with the emission of radiation according to Bohr's frequency condition.

14.9 Critical potentials

There are different methods of exciting the atoms. One of the most common method is to bombard the atoms with outside electrons having sufficient kinetic energies. This is achieved by accelerating the electrons by an electric field. Since the radiation of energy by the atoms during the de-excitation process takes place in a discontinuous manner as quanta of energy, it is to be expected that the absorption of energy involved in excitation and ionization of the atom take place in discrete quanta. This means that the potential difference of the accelerating electric field must have definite values to excite the atoms to its different quantized states. As long as the accelerating potential remains below a certain critical value, known as the *critical potential*, the colliding electron cannot excite the atom. Such collisions are known as elastic collisions where both the kinetic energy of translation and the momentum are conserved without any conversion of external kinetic energy into internal energy or vice versa. But if the accelerating potential exceeds the critical value mentioned above, the colliding electron loses a considerable amount of its kinetic energy and the atom struck by the electron suffers internal

changes of energy, which leads to its excitation or even ionization. Such collisions are called *inelastic collision*, to which the laws of classical impact can no longer be applied.

Critical potential : The least energy, expressed in electron volts, required to excite a free neutral atom from its ground state to a higher state is called the critical potential of the atom.

It is usual to distinguish two kinds of critical potentials, *viz.* excitation potential and ionization potential. *The excitation potential is the energy, expressed in electron volts, required to raise an atom from its normal state into an excited state.* It is also called *radiation potential*, because it causes the atom which has absorbed energy corresponding to that potential to emit radiations, when it returns from the excited state to the normal state. It is also known as *resonance potential*, because it is possible to provoke the same excitation of the atom and the subsequent emission of the same radiation as is caused by the bombarding electron.

The ionization potential of the atom may be defined as the energy required to knock an electron completely out of an atom and thereby ionize the atom.

Let us take a concrete example. We have seen that the energy of the n^{th} orbit of the hydrogen atom, expressed in electron volts, is given by $E_n = -\frac{13.6}{n^2}$. From this we get $-13.6, -3.4, -1.51, \ldots\ldots$ 0 eV for the energies of the 1st, 2nd, 3rd, ... ∞ orbits. Hence the energy required to raise the atom from the ground state ($n = 1$) to the first excited state is $13.6 - 3.4 = 10.2$ eV. The energy required to raise it to the second excited state is $(13.6 - 1.51) = 12.09$ eV and so on. The energy required to raise it to the ionized state is $(13.6 - 0) = 13.6$ eV. It is clear that 10.2, 12.09 volts are excitation potentials, while 13.6 volts is the ionization potential of the hydrogen atom.

Hydrogen atom has only one ionization potential but several excitation potentials. In the case of atoms with several electrons there can be more than one ionization potential. The energy required for the removal of outermost valence electrons, and hence least attracted by the nucleus compared to other electrons, is called the *first ionization potential.* Once the first electron is removed, the rest of the electrons are held more firmly on account of the relative increase of the positive

charge on the nucleus with respect to the electrons left behind. Hence the energy required to remove the second electron, *i.e.*, the *second ionization potential* will be higher than the first ionization potential and so on. Helium has two ionization potentials, 24.5 and 78.6 volts.

14.10 The Franck – Hertz Experiment

Convincing support to Bohr's prediction about the existence of discrete energy levels in atoms was provided by the study of the spectra of hydrogen like atoms. Further confirmations were provided by a series of entirely different type of experiments by J. Franck and G. Hertz in Germany within a short time after the formulation of Bohr's atom model.

The experimental arrangement is shown in Fig. 14.15. T is a sealed tube within which there is a metal filament F at one end which can be heated by passing an electric current through it so that electrons are thermionically emitted from F. At the other end of the tube there is a plate P which collects these electrons. G is a wire –

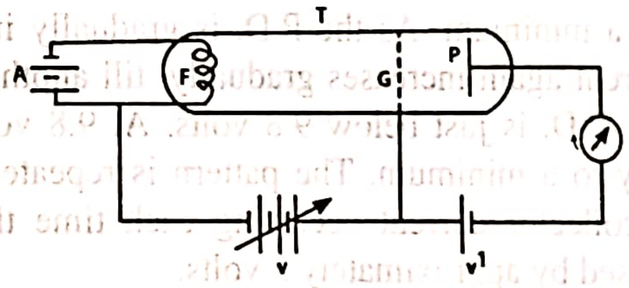


Fig. 14.15

grid placed just in front of P. Thus the electrons from F have to pass through this grid to reach P. The tube T contains a small amount of the experimental gas at a pressure of about 1 mm of mercury. Frank and Hertz used mercury vapour in their experiment.

The grid G is kept at a positive potential with respect to F so that the electrons are accelerated in traveling from F to G, their kinetic energy becoming maximum on reaching the grid G. The collector P is kept at a slightly negative potential with respect to G. This retarding potential V_0 is usually about 0.5 volt, so that if the electron energy is less than eV_0 while passing through G, they are unable to reach P.

Keeping V_0 constant, the potential difference V between F and G is gradually increased in small steps from zero upwards. A plot of collector current against the accelerating potential V is shown in

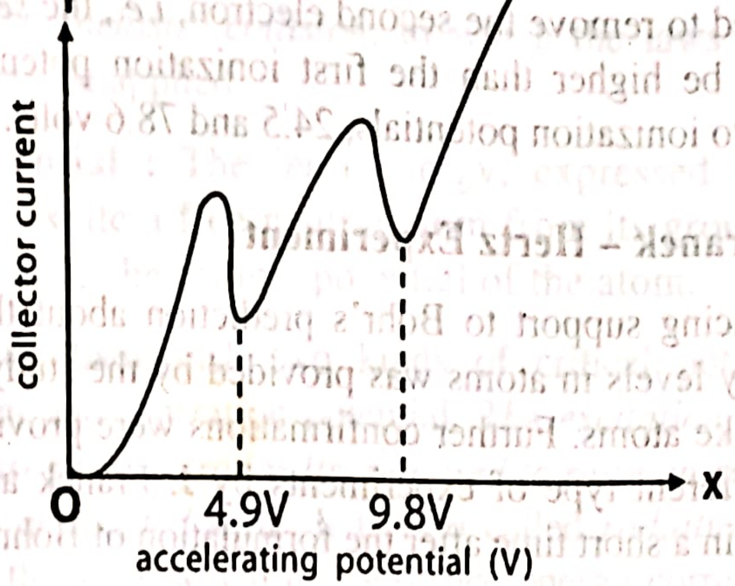


Fig. 14.16

Fig. 14.16. It can be seen from the graph that there is no collector current for V < 0.5 V. Above this, the collector current increases continuously. When the accelerating P.D. reaches a value of 4.9 volt, the current suddenly dips to a minimum. As the P.D. is gradually increased above 4.9 volts, the current again increases gradually, till another maximum is reached when the P.D. is just below 9.8 volts. At 9.8 volts, the current again dips steeply to a minimum. The pattern is repeated, a significant decrease in the collector current occurring each time the accelerating potential is increased by approximately 5 volts.

Explanation of the graph. An electron traveling from F to G collides with a gaseous atom transferring some of its energy to the atom. If the collision is elastic, then the energy transfer is negligibly small since the atoms are much heavier than the electrons. On the other hand if the collision is with an electron bound in the atom, then there may be considerable energy transfer from the impinging electron to the atomic electron.

According to Bohr atom model, an atomic electron can exist in different discrete energy levels. Normally it is in the ground level. If the energy of the impinging electron is not sufficient to raise the electron from the ground level to the first excited energy level by collision, then there will not be any energy transfer between the two inspite of the collision. Thus the electrons will reach G with their full energy and, if it is sufficient to overcome the retarding potential V_0, will fall on the collector plate P. The ammeter will then record a current. On increasing the accelerating potential, more and more electrons will reach P and the

collector current increases. Finally when the energy gained by the impinging electrons as they reach G becomes equal to the energy difference between the ground level and the first excited level of the atomic electrons, which is about 4.9 eV for mercury, then the entire energy of the incident electron is transferred to the atomic electron in an *elastic collision.* As a result the atomic electron is raised to the first excited level and the impinging electron is left with no energy to reach P. Consequently the collector current drops abruptly. However, the current does not reach zero because statistically some electrons may succeed in reaching the collector plate, avoiding an inelastic collision with a mercury atom. Such electrons record minimum current on the plate. Then as V is further increased, the electrons acquire more energy and start reaching P after the inelastic collisions, if they are left with energy to overcome V_0. Hence the plate current begins to rise again. As the electrons gain more and more energy with the increase of V, the collisions between the on-coming electrons and the atomic electrons now take place some distance behind G towards F. When the potential difference between F and G is twice the resonance potential (V=9.8 volts), the electrons coming from F gain energy equal to the energy difference between the two levels of the atom when it has traveled half way from F to G. So if it collides with an atom at this point, it causes transition of the atom from the ground level to the excited level and loses its entire energy in the process. It then starts afresh with zero kinetic energy towards G and again acquires the same amount of energy on reaching G. If it now collides with a second atom at this point, the atom will be raised to the upper level while the electron will lose its entire kinetic energy. As a result it will be unable to overcome the retarding potential so that current will suddenly drop again giving rise to the second peak. This explains the second minimum.

The other peaks arise in a similar manner as the electrons coming from F collide with larger number of atoms at different points along its path from F to G causing transitions between the levels of the corresponding atoms. These take place at potentials which are integral multiples of the resonance potentials.

Each time there is an inelastic collision, the electron in the ground state of mercury atom will be raised to the first excited state. The electron will soon return to the ground state by the emission of photons. By using spectroscopic techniques, the wavelength of the

radiation coming from the tube was found to be 2536 Å corresponding to the ground state. Let us calculate the energy of the photon corresponding to this wavelength

$$E = h\nu = \frac{hc}{\lambda} = \frac{(6.63\times10^{-34})(3\times10^{8})}{2536\times10^{-10}}\ \text{J}$$

$$= \frac{(6.63\times10^{-34})(3\times10^{8})}{(2536\times10^{-10})(1.6\times10^{-19})}\ \text{eV}$$

$$= 4.9\ \text{eV}.$$

Thus, the experiment shows that the energy lost by the electron in its elastic collision with the mercury atom reappears as a quantum of energy of wavelength $\frac{hc}{E}$. The experiment provides a convincing proof of the existence of discrete energy levels in the mercury atom.

14.11 Ionization Potential

The minimum energy required to set free an electron bound in the ground state of the atom is known as its ionization energy. This energy may be supplied to the atom by exposing it to photons or by bombarding it with electrons from outside. Fig. 14.17 shows the experimental arrangement for measuring the ionization energy of an atom. F is a metal filament within the sealed tube T which emits

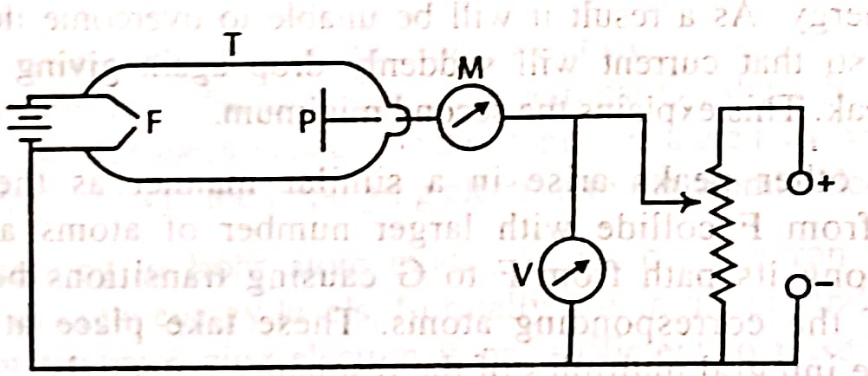


Fig. 14.17

electrons thermionically when heated - by passing an electric current through it. The emitted electrons are accelerated towards the plate P

which is maintained at a positive potential with respect to F. On reaching the plate P, the electrons give rise to a current which can be recorded by the ammeter M.

The tube T is maintained at a low pressure and contains a small amount of the experimental gas or vapour. As the potential difference between F and P is increased, the current I increases according to Child-Langmuir law ($I \propto V^{3/2}$). At a definite value of V = V_i, the rate of increase of current shows a sudden increase as shown in Fig. 14.18. The reason for this is that when $V > V_i$ the electron

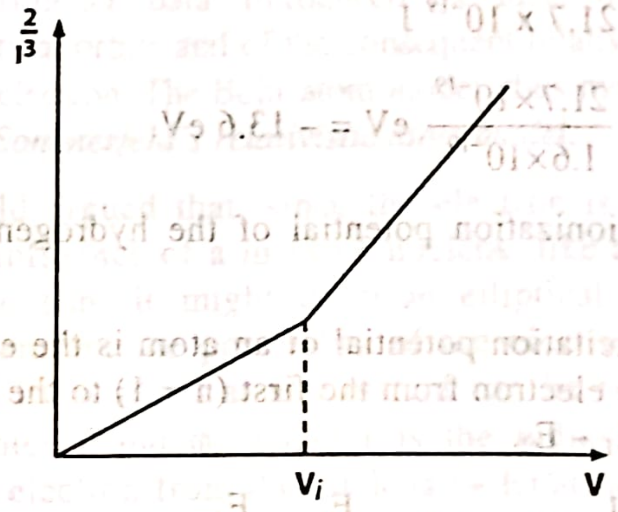


Fig. 14.18

coming from F acquire sufficient energy to ionize the gaseous atoms within T when they collide with these atoms. The electrons thus liberated from the atoms as a result of collisions, add to the swarm of electrons coming from F so that the current recorded by M begins to rise more rapidly with increasing V when the potential difference is greater than V_i.

The potential V_i is known as the *ionization potential.* An electron accelerated to this potential acquires energy equal to the ionization energy of the particular atom. The ionization potential and the resonance potential discussed in the Art. 14.9, are both referred to as the *critical potentials* of the atom.

***Example 14.12** Calculate (i) ionization potential and (ii) first excitation potential of the hydrogen atom. [h = 6.62 x 10^{-34} J s^{-1}, m = 9.1 x 10^{-31} kg and e = 1.6 x 10^{-19}C]*

Soln.

(i) The ionization potential is numerically equal to the energy of the electron in the first orbit.

$$E_1 = -\frac{me^4}{8\epsilon_0^2 h^2 n^2}$$

$$= -\frac{(9.1\times10^{-31})(1.6\times10^{-19})^4}{8(8.854\times10^{-12})^2(6.62\times10^{-34})^2} \quad (n = 1)$$

$$= -21.7 \times 10^{-19}\ \text{J}$$

$$= -\frac{21.7\times10^{-19}}{1.6\times10^{-19}}\ \text{eV} = -13.6\ \text{eV}.$$

Hence the ionization potential of the hydrogen atom is 13.6 volts.

(ii) The first excitation potential of an atom is the energy required to move the electron from the first (n = 1) to the second (n = 2) orbit, *i.e.*, $E_1 \sim E_2$.

$$E_n \propto \frac{1}{n^2}; \text{ Hence } E_2 = \frac{E_1}{2^2} = \frac{E_1}{4}$$

Now $E_1 = -13.6$ eV; $\quad \therefore \quad E_2 = -\frac{13.6}{4} = -3.4$ eV

$\therefore \quad E_1 \sim E_2 = 13.6 - 3.4 = 10.2$ eV.

14.12 Sommerfeld Relativistic Atom Model

Bohr's simple theory of circular orbits, in spite of its many successes, was found inadequate to explain certain details in the spectrum of hydrogen. According to Bohr, the lines in the hydrogen spectrum should each have a well-defined wavelength. Spectrographs of high resolving power showed that the H_α, H_β and H_γ lines in the hydrogen spectrum are not single. Each spectral line actually consisted of several very close lines packed together. Michelson found that under high resolution, the H_α line can be resolved into

two close components, with a wavelength separation of 0.13 A.U. This is called the *fine structure of the spectral lines.* This fine structure of spectral lines could not be explained by Bohr's theory which assumed that there was only one orbit for each quantum number n, whereas the observed fine structure suggested that for any given quantum number n there might be several orbits of slightly different energies. Moreover Bohr's theory applies to one-electron atoms and is not easily extended to describe more complicated or many electron atoms.

Sommerfeld, in 1915, in an attempt to fit Bohr's theory with the additional experimental data, introduced the ideas of motion of the electron in elliptical orbits and of the consequent relativistic variation of the mass of the electron. The Bohr atom model, thus modified, has come to be known as *Sommerfeld's relativistic atom model.*

Sommerfeld argued that, since the electron is moving around and under the influence of a massive nucleus, like a planet round a central massive sun, it might describe elliptical orbits as well. Considering therefore the possibility of the electron moving in an elliptical orbit, its position at any instant can be fixed in terms of polar co-ordinates r and φ, where r is the radius vector *i.e.*, the distance of the electron from the nucleus (+ E) at one of the foci of the ellipse and φ the vectorial angle *i.e.*, the angle which the radius vector makes with the major axis of the ellipse (Fig. 14.19). As the position of the electron is determined by two co-ordinates (r, φ), its motion is associated with two degrees of freedom and each of these two degrees of freedom will have to be separately quantized according

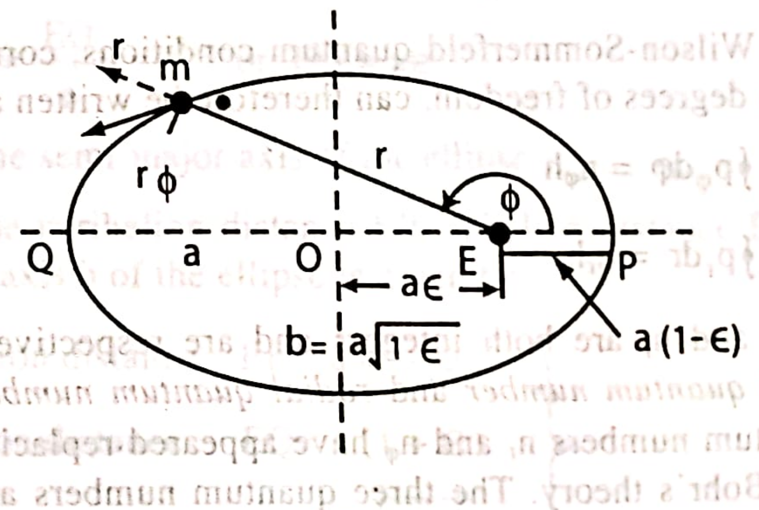


Fig. 14.19

to Wilson-Sommerfeld quantization rule. However, before writing the quantization conditions for the two degrees of freedom, the expressions for the potential and kinetic energies of the electron will be first written.

Assuming the potential energy of the electron to be zero when it is at infinite distance from the nucleus, its potential energy at a distance r from the nucleus will be given by

$$V = -\frac{eE}{4\pi\epsilon_0 r} = -\frac{e^2 Z}{4\pi\epsilon_0 r} \tag{14.55}$$

where e and E = Ze are the charges on the electron and the nucleus respectively.

The kinetic energy of the electron in its orbit will be given by

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(\dot{r}+r^2\dot{\varphi}^2)$$

$$= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 \tag{14.56}$$

where $\dot{r} = \frac{dr}{dt}$ is the radial velocity and $\dot{\phi} = \frac{d\phi}{dt}$ is the angular velocity of the electron.

Equations (14.55) and (14.56) involve two displacement co-ordinates r and φ and two momentum co-ordinates (angular and radial) given by,

$$p_\varphi = mr^2\dot{\varphi} \text{ and } p_r = m\dot{r} \tag{14.57}$$

The Wilson-Sommerfeld quantum conditions, corresponding to these two degrees of freedom, can therefore be written as

$$\oint p_\varphi d\varphi = n_\varphi h \tag{14.58}$$

$$\oint p_r dr = n_r h \tag{14.59}$$

where n_φ and n_r are both integers and are respectively known as *azimuthal quantum number* and *radial quantum number*. Thus two new quantum numbers n_r and n_φ have appeared replacing the single one n of Bohr's theory. The three quantum numbers are related by the equation

$$n = n_r + n_\varphi \qquad (14.60)$$

where n is called principal quantum number and is also known as the total quantum number. The principal quantum number n can take the integral values 1, 2, 3, ... etc. To determine the allowed elliptical orbits, the integrals in eqns. (14.58) and (14.59) must be evaluated. The momentum p_φ corresponding to the co-ordinate ϕ is merely the angular momentum p of the electron in the elliptic orbit and this from Kepler's law, is a constant, *i.e.*, $p_\varphi = p$ = constant. Integrating eqn. (14.58) from 0 to 2π,

$$\int_0^{2\pi} p_\varphi d\varphi = 2\pi p_\varphi = 2\pi m r^2 \dot{\varphi} = n_\varphi h$$

$$\text{or,} \quad p_\varphi = \frac{n_\varphi h}{2\pi} \qquad (14.61)$$

The condition represented by eqn. (14.61) is identical with Bohr's quantum condition for circular orbit.

The second integral, eqn. (14.59), can be evaluated following the method of Sommerfeld as shown below:

Let us consider the general equation of an ellipse in polar co-ordinates,

$$\frac{1}{r} = C_1 + C_2 \cos\varphi \qquad \text{(i)}$$

where C_1 and C_2 are constants to be determined for every ellipse. Let us then define the eccentricity $\in$ of the ellipse (Fig. 14.19) as the ratio EO/a *i.e.*,

$$\in = \frac{EO}{a} \quad \text{or, } EO = a\in$$

where a is the semi-major axis of the ellipse.

Now the perihelion distance EP, aphelion distance EQ and the semi-minor axis b of the ellipse is given by

$$\left.\begin{array}{ll} \text{Perihelion distance} & EP = a(1-\in) \\ \text{Aphelion distance} & EQ = a(1+\in) \\ \text{Semi-minor axis} & b = a\sqrt{1-\in^2} \end{array}\right\} \qquad \text{(ii)}$$

when the electron is at perihelion P, $\varphi = 0$ and $r = a(1 - \in)$ and eqn. (i) becomes

$$\frac{1}{a(1-\in)} = C_1 + C_2 \qquad \text{(iii.a)}$$

and when it is at the aphelion Q, $\varphi = \pi$ and $r = a(1 + \in)$ and eqn. (i) becomes

$$\frac{1}{a(1+\in)} = C_1 - C_2 \qquad \text{(iii.b)}$$

Adding (iii.a) and (iii.b) we get,

$$C_1 = \frac{1}{a(1-\in^2)}$$

and subtracting them we get,

$$C_2 = \frac{\in}{a(1-\in^2)}$$

Eqn. (i) of the ellipse, therefore, becomes

$$\frac{1}{r} = \frac{1+\in\cos\varphi}{a(1-\in^2)} \qquad \text{(iv)}$$

Taking logarithm of each side of eqn. (iv) and then differentiating with respect to φ we get,

$$\frac{1}{r}\frac{dr}{d\varphi} = \frac{\in\sin\varphi}{1+\in\cos\varphi} \qquad \text{(v)}$$

Now, the radial momentum,

$$p_r = m\dot{r} = m\frac{dr}{d\varphi}\cdot\frac{d\varphi}{dt} = m\dot{\varphi}\frac{dr}{d\varphi}$$

From eqn. (14.57)

$$m\dot{\varphi} = \frac{p_\varphi}{r^2}$$

$$\therefore \quad p_r = m\dot{r} = \frac{p_\varphi}{r^2}\left(\frac{dr}{d\varphi}\right) \qquad \text{(vi)}$$

Therefore the integrand of eqn. (14.59)

$$p_r dr = \frac{p_\varphi}{r^2}\left(\frac{dr}{d\varphi}\right) dr = \frac{p_\varphi}{r^2}\left(\frac{dr}{d\varphi}\right)\frac{dr}{d\varphi} \cdot d\varphi$$

$$= \frac{p_\varphi}{r^2}\left(\frac{dr}{d\varphi}\right)^2 d\varphi$$

Substituting the value of $\frac{1}{r}\left(\frac{dr}{d\varphi}\right)$ from eqn. (v)

$$p_r dr = p_\varphi \in^2 \frac{\sin^2\varphi}{(1+\in\cos\varphi)} 2d\varphi \qquad \text{(vii)}$$

The integral, eqn. (14.59), therefore becomes

$$\oint p_r dr = p_\varphi \int_0^{2\pi} \frac{\in^2 \sin^2\varphi}{(1+\in\cos\varphi)} d\varphi = n_r . h \qquad \text{(viii)}$$

Putting the value of p_φ from eqn. (14.61) as $\frac{n_\varphi h}{2\pi}$, we get

$$\frac{\in^2}{2\pi}\int_0^{2\pi} \frac{\sin^2\varphi}{(1+\in\cos\varphi)^2} d\varphi = \frac{n_r}{n_\varphi} \qquad \text{(ix)}$$

Putting $\int_0^{2\pi} \frac{\in^2 \sin^2\varphi}{(1+\in\cos\varphi)^2} d\varphi = I$

The integral $I = \int_0^{2\pi} \frac{\in^2 \sin^2\varphi}{(1+\in\cos\varphi)^2} d\varphi$ can be integrated by parts

$$\int u dv = uv - \int v du$$

Let $u = \in \sin\varphi$ so that $du = \in \cos\varphi \, d\varphi$

$$\text{and } dv = \frac{\in \sin^2\varphi}{(1+\in\cos\varphi)^2}\, d\varphi, \text{ so that}$$

$$v = \frac{1}{1+\in\cos\varphi}$$

Then

$$I = \left[\frac{\in\sin\varphi}{1+\in\cos\varphi}\right]_0^{2\pi} - \int_0^{2\pi} \frac{\in\cos\varphi}{1+\in\cos\varphi}\,.\,d\varphi$$

On substitution of the limits of integration, the first term on the R.H.S. becomes zero. Then

$$I = -\int_0^{2\pi} \frac{\in\cos\varphi}{1+\in\cos\varphi}\, d\varphi = \int_0^{2\pi}\left(\frac{1}{1+\in\cos\varphi} - 1\right)d\varphi$$

Now $\int_0^{2\pi} \frac{d\varphi}{1+\in\cos\varphi}$ is a standard integral whose value works out to be $2\pi/(1-\in^2)^{1/2}$.

Therefore,

$$I = \frac{2\pi}{(1-\in^2)^{1/2}} - 2\pi \qquad \text{(x)}$$

Now, eqn. (viii) can be written as

$$\frac{2\pi p_\varphi}{(1-\in^2)^{1/2}} - 2\pi p_\varphi = n_r h$$

$$\text{or,}\quad \frac{n_\varphi h}{(1-\in^2)^{1/2}} - n_\varphi h = n_r h \;\left(\text{since } p_\varphi = n_\varphi \,.\, \frac{h}{2\pi}\right) \qquad (14.62)$$

Here $\in$ is the eccentricity of the ellipse.

$$\therefore \quad n_r = \frac{n_\varphi}{(1-\in^2)^{1/2}} - n_\varphi$$

$$\text{or,}\quad n_r + n_\varphi = \frac{n_\varphi}{(1-\epsilon^2)^{1/2}}$$

But $n_r + n_\varphi = n$, the principal quantum number.

$$\text{Hence, } n = \frac{n_\varphi}{(1-\epsilon^2)^{1/2}}$$

$$\text{or,}\quad 1-\epsilon^2 = \frac{n_\varphi^2}{n^2} \qquad (14.63)$$

For an ellipse $1 - \epsilon^2 = b^2/a^2$ where a and b are the semi-major and semi-minor axes respectively.

Hence

$$\frac{b^2}{a^2} = \frac{n_\varphi^2}{n^2};\quad \text{or,}\quad \frac{b}{a} = \frac{n_\varphi}{n} \qquad (14.64)$$

Eqn. (14.64) is the condition that determines the allowed elliptical orbits. It indicates that out of all classically possible ellipses, only those ellipses are allowed by the quantum condition for which the ratio of major and minor axes is the same as that of the quantum numbers n and n_φ. When $n_\varphi = n$, $b = a$, $\epsilon = 0$ and the orbit becomes circular. n_φ cannot be zero, since the ellipse would then degenerate into a straight line passing through the nucleus. Also n_φ cannot be greater than n, since b is always less than a. Hence for a given value of n, n_φ can assume only n different values, which means there can be only n elliptical orbits of different eccentricities.

Example : ...... For the first orbit $n = 1$. Since $n_r + n_\varphi = 1$ and $n_\varphi \neq 0$, $n_\varphi = 1$. With $n = 2$, n_φ may have the values 1 or 2. So there are two possible orbits for $n = 2$, a circle and an ellipse. Similarly, there are three possible orbits for $n = 3$, a circle and two ellipses (Fig. 14.20). Usually the allowed orbits are described by giving values of n and n_φ. The three orbits for $n = 3$ are represented by symbols 3_3, 3_2 and 3_1,

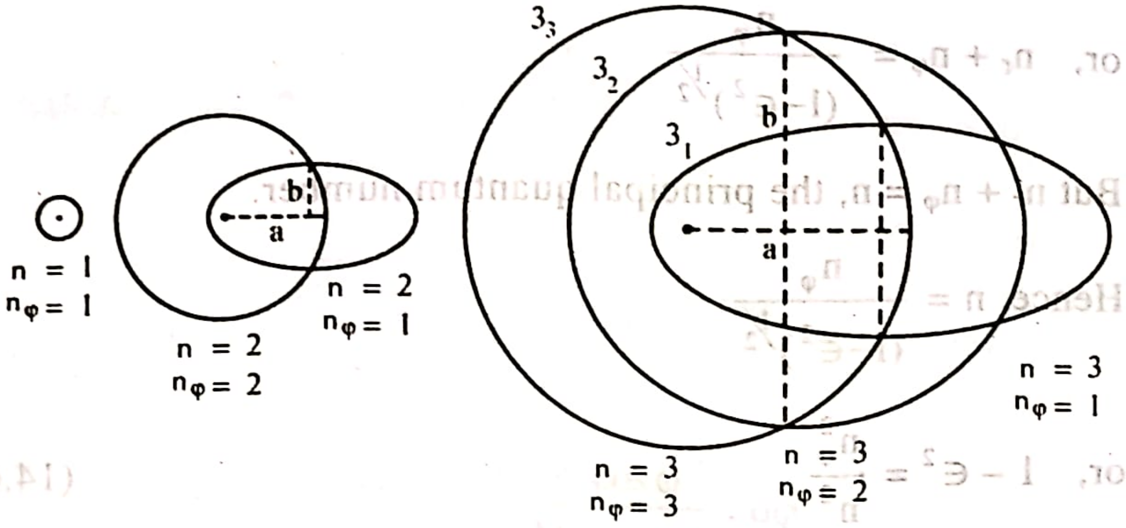


Fig. 14.20

the subscript being the azimuthal quantum number n_φ. In another notation, the azimuthal quantum number n_φ is described by the letters s, p, d, f, etc. The value of n_φ corresponding to these letters is 1, 2, 3, 4, etc. respectively. In this notation the orbit represented by 3s means that for this orbit n = 3 and $n_\varphi = 1$.

Total energy of stationary orbits :

The total energy of a single electron moving in an elliptical orbit is

$$E_n = P.\,E. + K.\,E.$$

The P. E. is $\dfrac{-Ze^2}{4\pi\epsilon_0\, r}$

The K. E. can be written as

$$\frac{1}{2}m\left[\left(\frac{dr}{dt}\right)^2 + \left(r.\frac{d\varphi}{dt}\right)^2\right]$$

where $\dfrac{dr}{dt}$ is the radial component of the velocity and $r\left(\dfrac{d\varphi}{dt}\right)$ is the transverse component of the velocity.

Hence,

$$E_n = \frac{-Ze^2}{4\pi\epsilon_0 r} + \frac{1}{2}m\left[\left(\frac{dr}{dt}\right)^2 + \left(r\frac{d\varphi}{dt}\right)^2\right]$$

$$= \frac{-Ze^2}{4\pi\epsilon_0 r} + \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\varphi}^2 \qquad (14.65)$$

On substitution from eqns. (14.57) and (vi), this equation becomes,

$$E_n = \frac{p_\varphi^2}{2m}\frac{1}{r^4}\left(\frac{dr}{d\varphi}\right)^2 + \frac{p_\varphi^2}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r}$$

$$= \frac{p_\varphi^2}{2mr^2}\left[\left(\frac{1}{r}\frac{dr}{d\varphi}\right)^2 + 1\right] - \frac{Ze^2}{4\pi\epsilon_0 r}$$

On substitution from eqns. (iv) and (v), the above eqn. becomes

$$E_n = \frac{p_\varphi^2}{2m}\frac{(1+\epsilon\cos\varphi)^2}{a^2(1-\epsilon^2)^2}\left[\frac{\epsilon^2\sin^2\varphi}{(1+\epsilon\cos\varphi)^2} + 1\right] - \frac{Ze^2(1+\epsilon\cos\varphi)}{4\pi\epsilon_0 a(1-\epsilon^2)}$$

$$= \frac{p_\varphi^2}{ma^2(1-\epsilon^2)^2}\left(\frac{1+\epsilon^2}{2} + \epsilon\cos\varphi\right) - \frac{Ze^2(1+\epsilon\cos\varphi)}{4\pi\epsilon_0 a(1-\epsilon^2)} \qquad (14.66)$$

Now the total energy for a conservative system is constant and is independent of time and φ. As $\cos\varphi$ varies during the motion of the electron, the total energy E_n will be constant only if the coefficients of $\epsilon\cos\varphi$ is zero. Therefore, collecting the coefficients of $\epsilon\cos\varphi$ and equating to zero,

$$\left[\frac{p_\varphi^2}{ma^2(1-\epsilon^2)^2} - \frac{e^2Z}{4\pi\epsilon_0 a(1-\epsilon^2)}\right] = 0$$

or, $$\frac{p_\varphi^2}{ma(1-\epsilon^2)} = \frac{e^2Z}{4\pi\epsilon_0}$$

$$\therefore \quad a = \frac{4\pi\epsilon\, p_\varphi^2}{me^2Z(1-\epsilon^2)} \qquad (14.67)$$

Substituting the value of p_φ and $(1 - \in^2)$ from eqns. (14.61) and (14.63), we get

$$a = 4\pi\in_0 \frac{n_\varphi^2 h^2}{4\pi^2} \cdot \frac{(n_r + n_\varphi)^2}{me^2 Z n_\varphi^2} = \frac{\in_0 h^2 (n_r + n_\varphi)^2}{\pi me^2 Z}$$

$$= \frac{\in_0 h^2 n^2}{\pi me^2 Z} \tag{14.68}$$

$\therefore$ Semi-major axis, $a = a_1 \frac{n^2}{Z}$ (14.69)

where $a_1 = \frac{\in_0 h^2}{\pi me^2}$ is the radius of the first Bohr circular orbit and the semi-minor axis,

$$b = a\sqrt{1-\in^2} = a_1 \frac{n^2}{Z}\frac{n_\varphi}{n} = a_1 \frac{n_\varphi n}{Z} \tag{14.70}$$

On substituting the value of 'a' as given by eqn. (14.67) into eqn. (14.66), the total energy is given by

$$E_n = \frac{p_\varphi^2 m^2 e^4 Z^2 (1-\in^2)^2}{16\pi^2 \in_0^2 m p_\varphi^4 (1-\in^2)^2}\left(\frac{1+\in^2}{2} + \in \cos\varphi\right)$$

$$- \frac{Ze^2 (1+\in\cos\varphi)(1-\in^2) me^2 Z}{16\pi^2 \in_0^2 p_\varphi^2 (1-\in^2)}$$

$$= \frac{me^4 Z^2}{16\pi^2 \in_0^2 p_\varphi^2}\left(\frac{1+\in^2}{2} - 1\right)$$

$$= \frac{me^4 Z^2}{32\pi^2 \in_0^2 p_\varphi^2}(\in^2 - 1)$$

$$= -\frac{Ze^2}{8\pi\in_0 a} \tag{14.71}$$

Substituting the value of 'a' from eqn. (14.68), eqn. (14.71) becomes

$$E_n = -\frac{me^4Z^2}{8\epsilon_0^2 n^2h^2} \tag{14.72}$$

The total energy given by eqn. (14.72) is exactly the same as that obtained by Bohr circular orbits and depends upon the reciprocal of the square of the principal quantum number n. The introduction of the elliptical orbits in place of circular orbits therefore does not introduce any new energy states for the hydrogen atom. However, as can be seen, the electron can now move in a number of orbits with the same energy. The relations (14.69) and (14.70) show that for a given value of n the major axis of the ellipse is independent of the azimuthal quantum number n_φ, the minor axis does depend upon n_φ. There will consequently be different ellipses for different values of n_φ. As already pointed out, for n = 3, n_φ can have values 1, 2, 3. The possibility of n_φ being equal to zero is ruled out as then the ellipse will reduce to a straight line with the nucleus at one end. The energy state with $n_\varphi = 0$ was therefore excluded by Sommerfeld on the ground that the electron will collide with the nucleus. Thus for n = 3, there are three quantized orbits corresponding to $n_\varphi = 1, 2, 3$. All the three ellipses will have the same major axis, but their minor axes are given by

for $n_\varphi = 1$, $2b = \frac{2a}{3}$

for $n_\varphi = 2$, $2b = \frac{4a}{3}$

and for $n_\varphi = 3$, $2b = \frac{6a}{3}$

or, for $n_\varphi = 3$, b = a *i.e.*, the ellipse will be a circle.

The multiplicity of orbits of different eccentricities (due to different values of b) does not introduce new energy states as mentioned above, and therefore there will be no new energy transitions and corresponding lines in the spectrum.

Sommerfeld's relativistic correction for variation of electron mass :

The spectral lines emitted by even hydrogen atom, the simplest of all atoms, have fine structure. For example, both H_α and H_β were found to be not single lines, but close doublets with separation of 0.14 and 0.08 Å respectively by Michelson with his interferometer. This was confirmed by many subsequent experiments. Sommerfeld's extension of Bohr's theory introducing elliptic orbits did not, however, introduce any new energy states to the hydrogen atom to explain the existence of fine structure. For a given total quantum number n, all elliptical orbits with different values of n_φ have just the same energy as the Bohr circular orbit with the same n.

It was suggested by Bohr that although the velocity of the electron in a circular orbit is constant, it varies in an elliptical orbit, being maximum at the perihelion and minimum at the aphelion. Furthermore, this velocity is quite large $(\simeq \frac{1}{137}c)$. Thus the relativistic change in the mass of the orbital electron should be taken into account in computing the energy levels. Following this suggestion, Sommerfeld applied special theory of relativity to calculate the new energy levels, taking into account the variation of the mass of the electron with velocity. As the electrons in orbits having the same total quantum number n but different azimuthal quantum number n_φ will have different velocities, the mass of the electrons and hence the resultant energy levels will all be different. Sommerfeld showed that the relativistic equation describing the path of an electron is

$$\frac{1}{r} = \frac{1+\epsilon\cos(\psi\varphi)}{a(1-\epsilon^2)} \tag{14.73}$$

$$\text{where } \psi^2 = 1 - \frac{Z^2e^4}{16\pi^2\epsilon_0^2 p^2c^2}$$

The path of the electron given by equation (14.73) is a *rosette* – a *precessing ellipse i.e.*, an ellipse whose major axis precesses slowly in the plane of the ellipse about an axis through the nucleus (Fig. 14.21).

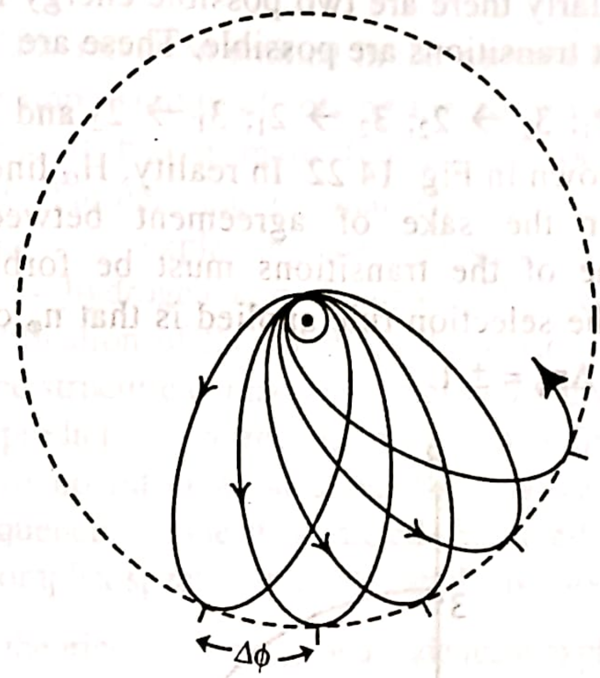


Fig. 14.21

It can be shown that the expression for the total energy in the relativistic model is given by

$$E_n = \frac{me^4Z^2}{8\epsilon_0^2 h^2 n^2} - \frac{me^4Z^4\alpha^2}{8\epsilon_0^2 h^2}\left(\frac{n}{n_\varphi} - \frac{3}{4}\right)\frac{1}{n^4} \tag{14.74}$$

where $\alpha = \frac{e^2}{2\epsilon_0 ch} \simeq \frac{1}{137}$

α is a dimensionless quantity and is called the fine *structure constant*. The first term in eqn. (14.74) is the energy of the electron in the orbit with the principal quantum number n. The second term is *Sommerfeld's relativistic* correction. The second term shows that the energy does depend on the azimuthal quantum number n_φ. This results in the splitting of the energy levels of the atom into a number of components. For a given value of n, the number of components is n corresponding to n permitted values of n_φ.

Fine structure of the H_α line

H_α line arises due to transition of the electron from n = 3 state to n = 2 state of the hydrogen atom. For n = 3, there are three possible energy levels corresponding to the n permitted values of n_φ,

1, 2, and 3. Similarly there are two possible energy levels for n = 2. Theoretically, six transitions are possible. These are

$3_3 \rightarrow 2_2$; $3_3 \rightarrow 2_1$; $3_2 \rightarrow 2_2$; $3_2 \rightarrow 2_1$; $3_1 \rightarrow 2_2$ and $3_1 \rightarrow 2_1$. These transitions are shown in Fig. 14.22. In reality, H_α line has only *three* components. For the sake of agreement between theory and experiment, some of the transitions must be forbidden by some *selection rule.* The selection rule applied is that n_φ can change only by +1 or – 1 *i.e.*, $\Delta n_\varphi = \pm 1$.

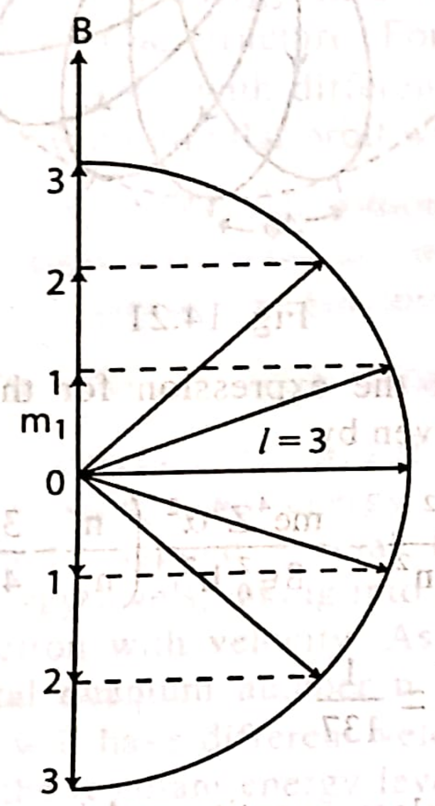


Fig. 14.22

Δn, however, has no such restriction. Therefore, on the basis of this selection rule, there can be only three allowed transitions. These are shown by solid lines in the figure while the forbidden transitions are shown by broken lines.

14.13 The Vector Atom Model

Bohr's theory was able to explain only the series spectra of the simplest atom, hydrogen or hydrogen like. In the case of atoms having a swarm of 30, 40 or even more electrons whirling round the nucleus, it seemed impossible to calculate the energy of the system and from this

the frequency of the radiation emitted in the passage between two states E_{n1} and E_{n2}. The impossibility of solving the problem was not merely a mathematical one. It had a more fundamental origin in the basic conception itself. Further, the original simple theory of Bohr was absolutely incapable of explaining the fine structure of spectral lines even in the simplest hydrogen atom. Sommerfeld's theory, though giving a theoretical justification of the splitting of individual spectral lines of hydrogen into fine structure components, met with only a partial success, as it could not predict the correct number of the fine structure lines. Moreover, it gave no information about the relative intensities of the lines, whose frequencies alone it predicted. Sommerfeld's theory could not explain the complex spectra of alkali metals like sodium.

These older theories could not give an adequate explanation of the new discoveries like Zeeman effect, Stark effect, in which the spectral lines could be split up under the influence of magnetic and electric fields. This failure brought out the insufficiency of the older ideas with still greater force.

Another drawback of the Bohr atom model was that it could not explain how the orbital electrons were distributed in atoms.

However, the most fundamental objection to be raised against Bohr atom model was that it involved the use of two theories which are essentially opposed to each other. The quantum theory was invoked to account for the existence of stationary orbits and for the frequencies of the radiation emitted, while the motion of the electrons in the orbits was controlled by the laws of classical mechanics.

New ideas were therefore advanced in order to explain the complex nature of atomic spectra and their relation to atomic structure, which finally resulted in what is now known as the *vector atom model*. Among the physicists who made substantial contribution to this model, mention must be made of Bohr, Sommerfeld, Uhlenbeck, Goudsmit, Pauli, Lande, Stern and Gerlach.

The two essential elements that characterize the vector atom model and differentiate it from the other models are (i) spatial quantization or quantization of direction and (ii) the spinning electron hypothesis.

Spatial Quantization: According to Bohr-Sommerfeld atom model, the orbits are quantized as regards their magnitude *i.e.*, their size and form only. But quantum theory demands more than this, *viz.*, they

should be quantized as regards their direction also, *i.e.*, out of the continuous manifold of all possible positions of the orbits in space, permitted by classical mechanics, only a discrete number of positions conformable to quantum conditions should be selected. This means that the orbits will be *vector* quantities.

To quantize spatially, a certain preferred direction with respect to which the orbits may receive their orientation is, however, needed. This preferred direction is chosen as the direction of an external magnetic field that is applied to the atom. The different permitted orientations of an electron orbit are determined by the fact that *the projections of the quantized orbits on the field direction must themselves be quantized.*

This new idea of spatial quantization was first proposed and worked out on the basis of the quantum theory by Sommerfeld in connection with the interpretation of the Zeeman effect in which the spectral line splits up under the influence of an external magnetic field.

Spinning electron: To account for the observed fine structure of spectral lines and to explain the anomalous Zeeman effect, two Dutch physicists, G.E. Uhlenbeck and S. A. Goudsmit, put forward in 1925 the hypothesis of the spinning electron. According to this hypothesis, the electron revolves not only in an orbit round the nucleus but also about an axis of its own, somewhat like a planet in the solar system. In other words, the electron is endowed with a spin motion over and above the orbital motion.

The spinning electron hypothesis brought in profound modifications in the atom model. In general, a body rotating about an axis gives rise to a mechanical angular momentum. Further, if the rotating body carries an electric charge, the latter will also revolve along with the body. But a revolving charge is equivalent to a circuital current, giving rise to a magnetic moment. Hence the rotation of a charged body about an axis produces a mechanical momentum as well as a magnetic moment.

In the older atom model, the electron was supposed to have only orbital motion round the nucleus, which would produce a mechanical momentum and magnetic moment. This new twist, *i.e.*, the spinning electron hypothesis endowes the electron with two angular momenta and two magnetic moments, one due to orbital motion and the other due to spin. In consequence, the total angular momentum of the atom will no longer be due simply to the orbital motion of the electron, but also due to

the spin of the electron. Like wise, the total magnetic moment of the atom is due to both the orbital and spin magnetic moments.

According to quantum theory, the spin motion of the electron, like its orbital motion, should be quantized. Hence a new quantum number, called the *spin quantum number*, is to be introduced in addition to the orbital quantum number. Since the orbital and spin motions are both quantized in magnitude and direction according to the idea of spatial quantization, they are considered as *quantized vectors*. Since the different component parts that determine the state of the atom, such as orbital and spin motions, are all quantized vectors, the atom model built on such considerations, is aptly called the *vector atom model*, to which vector laws apply.

Quantum number associated with the vector atom model.

In the vector atom model, a quantum number is assigned to each of the component parts. The numerical value of the quantum number may conveniently be thought of as the length of the vector which represents the angular momentum of that component part. In vector analysis, angular momentum is represented by a straight line whose direction is parallel to the axis of rotation and whose length is proportional to the magnitude of the momentum.

(1) The principal or total quantum number (n) : This is identical with the one used in Bohr-Sommerfeld theory and represents the serial number of the shells starting from the innermost shell. It can take only integral values excluding zero, *i.e.*, n = 1, 2, 3, 4, ...

(2) The orbital quantum number (ℓ): This may take any values between 0 and (n – 1), both inclusive. Thus if n = 4, ℓ can have values 0, 1, 2, 3. The orbital angular momentum p_ℓ of the electron is given by $p_\ell = \frac{\ell h}{2\pi}$. It may, however, be noted that according to wave mechanics, the orbital angular momentum has the value, $p_\ell = \sqrt{\ell(\ell+1)}\,\frac{h}{2\pi}$. For example, an electron whose orbital quantum number is 2, has the angular momentum

$$p_\ell = \sqrt{2(2+1)}\,\frac{h}{2\pi} = 2.6 \times 10^{-34} \text{ J.s.}$$

By convention, an electron for which $\ell = 0$ is called an *s* electron; if $\ell = 1$, *p* electron; $\ell = 2$, *d* electron; $\ell = 3$, f electron; $\ell = 4$, *g* electron, etc.

(3) The spin quantum number (s): The spin quantum number has only one magnitude, $s = \frac{1}{2}$. The spin angular momentum number, $p_s = s.\frac{h}{2\pi}$ where $s = ½$. According to wave mechanics, $p_s = \sqrt{s(s+1)}\ \frac{h}{2\pi}$.

(4) Total angular quantum number (j): It refers to the resultant angular momentum of the electron due to both orbital and spin motions and is the numerical value of the vector sum of ℓ and *s*. The vector $\vec{j}$ is, therefore, defined by the equation $\vec{j} = \vec{\ell} + \vec{s}$, with the restriction that $\vec{j}$ is positive. The value of $\vec{j}$ is naturally a half-integer, since one of its components *s* is always equal to ½. It is usually expressed as $\vec{j} = \vec{\ell} \pm \vec{s}\ (\vec{\ell} \pm \frac{1}{2})$, *plus* sign when *s* is *parallel to* ℓ and *minus* sign when *s* is *antiparallel*. Thus for $\ell = 3$, j can have the values $\frac{7}{2}$ $(\ell + \frac{1}{2})$ and $\frac{5}{2}$ $(\ell - \frac{1}{2})$. The total angular momentum of the electron, $p_j = \frac{jh}{2\pi}$ and more correctly by $\sqrt{j(j+1)}\ \frac{h}{2\pi}$ according to wave mechanics.

Magnetic quantum numbers: To explain the splitting of spectral lines in a magnetic field, three more magnetic quantum numbers are introduced. These are

(5) Magnetic orbital quantum number (m_ℓ): The projection of the orbital quantum number ℓ on the magnetic field direction is called the magnetic orbital quantum number m_ℓ. m_ℓ is an integer and may have $(2\ell+1)$ values from $-\ell$ to $+\ell$ including zero. This is because, according to spatial quantization, the projection of ℓ in the field direction must itself be quantized. Hence ℓ can be inclined to the field direction only at such discrete angles that its projection m_ℓ may also be an integer. For example, if the vector ℓ is inclined to the field direction at an angle θ, its projection $m_\ell = \ell \cos\theta$. Now since m_ℓ has to be an integer and cos θ cannot exceed unity, the

permitted values of m_ℓ are from $+\ell$ to $-\ell$ at unit intervals, *viz.* ℓ, $(\ell - 1)$, $(\ell - 2)$, ... 1, 0, – 1, ... $-(\ell - 1)$, $-\ell$. Hence the total number of possible values of m_ℓ is $(2l + 1)$. Thus for $\ell = 2$, the permitted orientations of ℓ are 5 for which $m_\ell = +2, +1, 0, -1, -2$ as shown in Fig. 14.22.

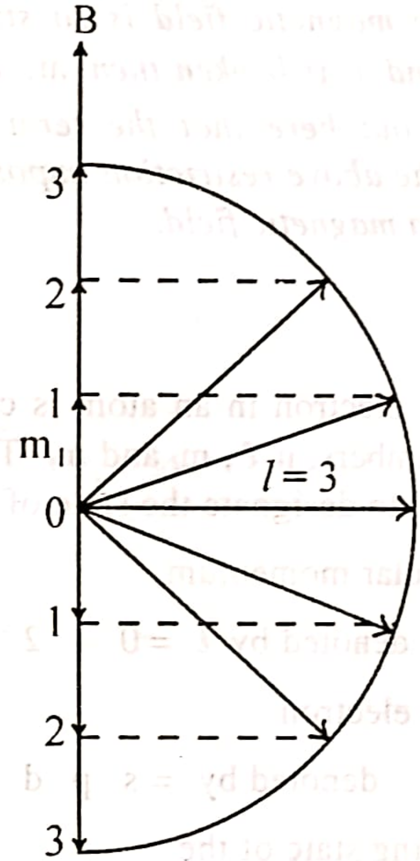


Fig. 14.22

(6) Magnetic spin quantum number (m_s): This is the numerical value of the projection of the spin vector s on the field direction. By analogy with the orbital vector ℓ, the vector s can have also only (2s + 1) permitted orientations with respect to the field direction. Consequently m_s can have any of the (2s + 1) values from –s to +s, without, however, including zero, since s is always equal to ½ and never zero. Thus m_s can have only two values +½ or – ½.

(7) Magnetic total angular momentum quantum number (m_j): This is the projection of the total angular momentum vector j on the direction of the magnetic field. Since we are dealing with single

electron, j can have only half integral values, since $j = \ell \pm \frac{1}{2}$. Consequently m_j must have only half-integral values. As the permitted orientations of j with respect to the field direction are (2j + 1), m_j can have only (2j + 1) values, from + j to – j, zero excluded.

Note : *The quantum number m_j is effective in ordinary magnetic fields, but when the magnetic field is so strong that the coupling between vectors s and ℓ is broken then m_ℓ and m_s come into play. It may be pointed out here that the term space quantization is usually applied to the above restriction imposed on the vectors, ℓ, s, j in the presence of a magnetic field.*

System of notation

The state of an electron in an atom is completely specified by the four quantum numbers, n, ℓ, m_l and m_s. The following system of notation is employed to designate the state of an electron or atom.

The orbital angular momentum of electron denoted by ℓ	= 0 1 2 3 4 5 6 7
The state of the electron denoted by	= s p d f g h i j
The corresponding state of the atom denoted by	= S P D F G H I J

The state of an electron is more completely specified by putting the principal quantum number n before the letter that specifies the value of the vector ℓ for the orbit. Thus a 2s electron is one for which n = 2 and ℓ = 0. The corresponding states of the atom are designated by capital letters, but provided only a single electron is involved in energy level changes (*e.g.*, in the case of hydrogen and alkali metals). Thus, an electron in 2p state will be an atom in a 2P state.

Further, for each value of ℓ, other than zero, there are two values of j given by $\ell + \frac{1}{2}$ and $\ell - \frac{1}{2}$. The two values correspond to two energy level terms.

Therefore, an electron, formerly considered to be in an energy level denoted by 2p ($n = 2, \ell = 1$), can actually exist in two energy levels close together. This gives rise to the observed doublet in the spectrum. Thus, in the case of an electron in the 2p state.

$$n = 2,\ \ell = 1,\ \ j = 1 - \tfrac{1}{2} = \tfrac{1}{2}$$

$$\text{or,}\ \ n = 2,\ \ell = 1,\ \ j = 1 + \tfrac{1}{2} = \tfrac{3}{2}$$

Thus the 2p electron is designated either by $2p_{1/2}$ or $2p_{3/2}$. The abbreviated notation for these states is $2^2p_{1/2}$ denoting the multiplicity of the level as 2.

The s level is single but for uniformity 2 is also written before s, *e.g.*, $3^2s_{1/2}$.

The transition between various energy levels of the atom are restricted by the selection rules that ℓ can only change by $\pm$ 1, *i.e.*, $\Delta \ell = \pm 1$ while j can change by 0 or $\pm$ 1, *i.e.*, $\Delta j = 0$ or $\pm$ 1.

The wavelengths 5890 A and 5896 A of the well-known sodium doublet can be explained as arising from the transition of the atomic states

$$3\,P_{1/2} \longrightarrow 3\,S_{1/2} \qquad (5896\ \text{A})$$

$$\text{and}\ 3\,P_{3/2} \longrightarrow 3\,S_{1/2} \qquad (5890\ \text{A})$$

Briefly, both lines can be expressed by

$$3^2\,P_{1/2} \longrightarrow 3^2\,S_{1/2}.$$

Coupling schemes

In an atom having two or more electrons, there are several ways in which the different vectors of the electrons may combine to give the vectors representing the atom as a whole. The method of combination depends on the interaction or coupling between the component vectors, since the orbital and spin motions of the electron producing magnetic fields result in mutual perturbation.

Two types of combination have been developed. They are (i) *Russell-Saunders* or *L – S coupling* and (ii) *j – j coupling*.

(1) L – S coupling. In this type of coupling all the orbital momentum vectors ℓ of the various electrons combine to form a resultant **L** and all the spin angular momentum vectors s likewise combine independently to form a resultant **S**. The total angular momentum of the atom is then given by

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

This scheme may be summarized as follows $L = \sum \ell_i$; $S = \sum s_i$; $J = L + S$. L is always an integer including zero. S is an integer for an even number of electrons, and odd multiple of $\frac{1}{2}$ for an odd number of electrons. Hence J must be an integer if S is an integer and J must be an odd multiple of $\frac{1}{2}$ if S is an odd multiple of $\frac{1}{2}$. The L – S coupling is illustrated in Fig. 14.23 for L = 2, S = 1 and L = 2, S = $\frac{3}{2}$.

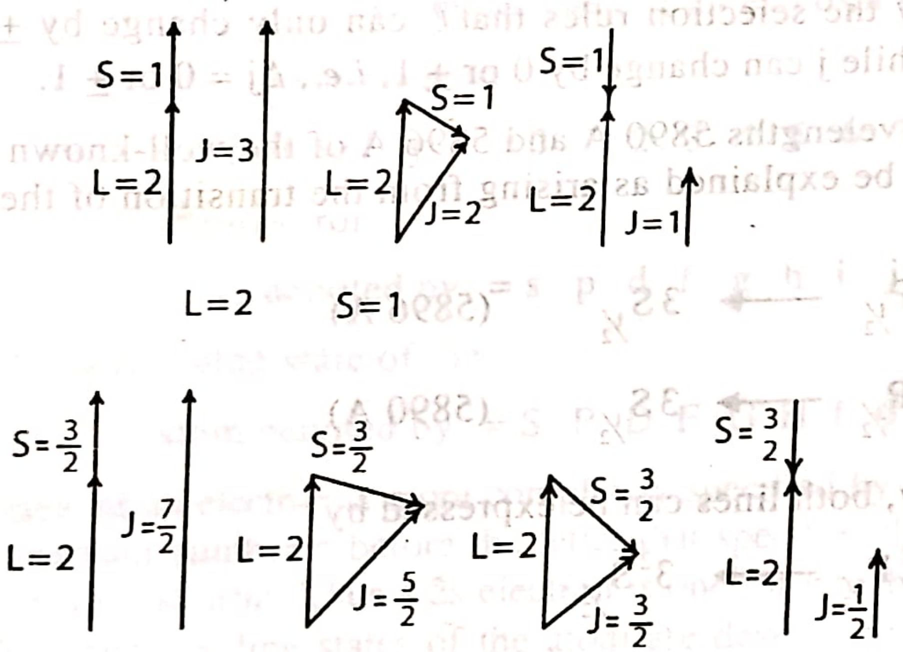


Fig. 14.23

It can be shown that when L > S, J can assume (2S + 1) values and when L < S, it can assume (2L + 1) values. If L = 0, J can have only one value S = J.

(2) The J – J coupling. Under certain circumstances, the interaction between the spin and the orbital vectors of each electron

is stronger than the interaction between either the spin vectors or the orbital vectors of the different electrons. If this is the case, each electron is considered separately and its contribution to the total angular momentum is obtained by first combining its individual spin and orbital vectors by the relation $j = l + s$. The vector sum of all the individual j vectors of the electrons gives the total angular momentum J of the atom. Thus

$$J = \sum j_i \quad \text{where } j = \ell + s$$

L – S coupling and J-J coupling are limiting cases, between which a whole range of intermediate types may occur, which makes the problem very difficult for treatment. For most known cases, however, the L – S coupling is effective. J-J coupling exists mainly in heavy atoms.

Application of spatial quantization

According to quantum theory, the resultant vectors L, S and J, representing the atom and obtained by the coupling mechanisms mentioned above, are quantized in magnitude and direction. Hence the number of permitted orientations that L, S and J can have with respect to a given field direction are (2L + 1), (2S + 1) and (2J + 1) respectively. The corresponding magnetic quantum numbers

$$m_L = \sum m_\ell \,, \; m_S = \sum m_s \text{ and } m_J = \sum m_j$$

can only have (2L + 1), (2S + 1) and (2J + 1) values respectively.

It may be noted that in the one electron system, *i.e.*, *in atom with a single effective electron, the state of atom as a whole is identical with the state of the electron*, so that $L = \ell$, $S = s$ and $J = j$.

14.14 The Pauli's Exclusion Principle

Out of the quantum numbers n, ℓ, m_ℓ, s, m_s, j and m_j discussed in Art. 14.13, a set of only four is sufficient for completely defining the quantum state of an electron in an atom. For example – if an atom is placed in a magnetic field strong enough to break the coupling between ℓ and s, the quantum numbers n, ℓ, m_ℓ and m_s completely define the quantum state of an electron in the atom. In

ordinary magnetic field or in the absence of the magnetic field, the quantum numbers n, ℓ, j and m_j are sufficient to completely define the quantum state of an electron in the atom.

When one deals with atoms with two or more electrons, the question that arises is how the quantum numbers should be assigned to these electrons in order to completely define their respective quantum states. The rule that solves this problem was proposed by Wolfgang Pauli. Based on a study of the transitions that are present, and those that are expected but not present, in the emission spectra of atoms, Pauli in 1925, put forward a principle, known as *Pauli's exclusion principle*. Simply stated, the Pauli's exclusion principle is as follows:

No two electrons in a single atom can have the same set of quantum numbers (n, ℓ, m_ℓ, m_s).

The principle implies that each electron in an atom must have a different set of quantum numbers n, ℓ, m_ℓ, m_s. If two electrons have exactly identical set of quantum numbers, then one of those electrons would be excluded from entering into the constitution of the atom. Hence the name exclusion principle.

Application of Pauli's exclusion principle

The number of electrons that can occupy a given sub-shell of an atom can be calculated with the help of this principle.

(1) Let us consider the K-shell. For this shell n = 1 and $\ell = 0$. since $\ell = 0$, m_ℓ is equal to zero and since $s = \frac{1}{2}$, m_s can be either +½ and –½. Hence the K-shell can have only two electrons – one with quantum numbers n = 1, $\ell = 0$, $m_\ell = 0$ and $m_s = +\frac{1}{2}$ and the other with quantum numbers n = 1, ℓ = 0, $m_\ell = 0$ and $m_s = -\frac{1}{2}$. If there were a third electron, its quantum numbers must be identical with those of either of these two electrons. This is not allowed according to Pauli's exclusion principle. The K-shell is therefore completed or closed with two electrons.

(2) For the L-shell n = 2 and ℓ = 0 or 1. For the sub-shell n = 2 and ℓ = 0, m_ℓ = 0 and $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$. Hence there can be only 2 electrons in this sub-shell. For the sub-shell n = 2 and ℓ = 1, m_ℓ can have values of – 1, 0, +1. For each of these three values of m_ℓ, m_s may be either $+\frac{1}{2}$ or $-\frac{1}{2}$. Hence there will be six possible set of values for the quantum numbers characterizing the electrons. Therefore, the maximum number of electrons that can be accommodated in this sub-shell is 6. The L-shell with two sub-shells [(n = 2, ℓ = 0) and (n = 2, ℓ = 1)] can, therefore, accommodate a maximum number of 2 + 6 = 8 electrons. In other words, the L-shell will be complete when it contains 2 + 6 = 8 electrons.

(3) For the M-shell, n = 3. Hence there will be three sub-shells with ℓ = 0, 1, 2. As explained above, the first and second sub-shells are completed with 2 and 6 electrons respectively. For the third sub-shell, ℓ = 2. Hence m_ℓ will have (2ℓ + 1) values [$-2\ell, -\ell, 0, \ell, 2\ell$] and for each value of m_ℓ, m_s can be $+\frac{1}{2}$ or $-\frac{1}{2}$. Hence the third sub-shell is completed with 2(2ℓ+1) *i.e.*, 10 electrons since ℓ = 2. Hence the total number of electrons required to complete the M-shell is 2+6+10 = 18.

(4) Similarly the N-shell (n = 4) with four sub-shells can have a maximum of 32 electrons.

In general, the two following conclusions can be made:

(i) In the n^{th} shell there are n sub-shells corresponding to the values 0, 1, 2, 3, ... (n – 1) of ℓ.

The maximum number of electrons in a sub-shell with a given value of ℓ is 2 (2ℓ+1).

Orbital quantum number (ℓ)	0	1	2	3	4	......
No. of possible electron states	2	6	10	14	18	......
Sub-shell symbol	s	p	d	f	g	......

(ii) The number of electrons that can be accommodated in a shell with principal quantum number n = sum of the electrons in the constituent n sub-shells.

$$= \sum_{\ell=0}^{\ell=n-1} 2(2\ell+1) = 2\sum_{\ell=0}^{\ell=n-1}(2\ell+1)$$

$$= 2\,[1 + 3 + 5 + 7 + \ldots\ldots\ldots\{2\,(n-1)+1\}]$$

$$= 2n^2.$$

The distribution of electrons in different shells according to this scheme is shown in the following table

Shell symbol	K	L	M	N	O
Quantum number (n)	1	2	3	4	5
Number of electrons ($2n^2$)	2	8	18	32	50

The, distribution of electrons in the various shells and sub-shells according to the exclusion principle is given in the following table.

n	ℓ	m_ℓ	m_s	No. of electrons in sub-shell with spectroscopic notation		Total no. of electrons in shells = $2n^2$
1	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$1s^2$	2
2	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$2s^2$	
2	1	−1, 0, + 1	$+\frac{1}{2}, -\frac{1}{2}$	6	$2p^6$	18
3	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$3s^2$	
3	1	−1, 0, +1	$+\frac{1}{2}, -\frac{1}{2}$	6	$3s^6$	18
3	0	−2, −1, 0, +1, +2	$+\frac{1}{2}, -\frac{1}{2}$	10	$3d^{10}$	
4	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$4s^2$	
4	1	−1, 0, +1	$+\frac{1}{2}, -\frac{1}{2}$	6	$4p^6$	
4	2	−2, −1, 0, +1, +2	$+\frac{1}{2}, -\frac{1}{2}$	10	$4d^{10}$	32
4	3	−3, −2, −1, 0, +1, +2, +3	$+\frac{1}{2}, -\frac{1}{2}$	14	$4f^{14}$	

14.15 The Periodic Table

The periodic table is an arrangement of different elements that exist in nature, based on their chemical properties and atomic numbers. The elements are arranged in seven horizontal rows, called *periods*, and eight vertical columns called *groups*. The elements arranged in rows have their chemical and physical properties varying gradually as a periodic function of atomic number and the elements in a group exhibit similar properties. Since atomic number of an element gives also the number of electrons in the atom, it follows that the atoms of successive elements in the periodic table are formed by the addition of one more electron at each step. Let us apply Pauli's exclusion principle for determining the distribution of electrons in different shells and sub-shells in the atoms of these elements. We are interested here in the normal state of the atom in which the electrons are in their lowest possible energy levels. The resulting electron configurations of a few elements are given below:

(1) Hydrogen (Z = 1)

Hydrogen is the first element in the periodic table. The quantum numbers for the lowest energy state of the hydrogen atom are $n = 1$, $\ell = 0$, $m_\ell = 0$ and $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$. The only electron of hydrogen which is 1s electron can have a spin magnetic quantum number m_s equal to either $+\frac{1}{2}$ or $-\frac{1}{2}$.

(2). Helium (Z = 2)

The next element in the periodic table is helium which has two 1s electrons with quantum numbers $n = 1$, $\ell = 0$, $m_\ell = 0$ and $m_s = +\frac{1}{2}$ for one electron and $-\frac{1}{2}$ for the other electron. The symbolic representation is $1s^2$. The K-shell is completely filled or closed. The rectangular enclosure indicates that the electrons are interlocked in a very stable arrangement of electrons; helium is thus an *inert gas*.

(3) Lithium (Z = 3)

It has three electrons. Two electrons can be put into the shell $n = 1$, $\ell = 0$. The third electron must be put into a new shell $n = 2$, $\ell = 0$. The neutral lithium atom is, therefore, represented by $1s^2 2s$. Lithium is one of the alkali elements and has a valence of unity. The transition from the first quantum orbit ($n = 1$) to the second quantum

orbit (n = 2) is accompanied by a great increase in the orbital dimensions. According to calculation based on simple hydrogen atom model, this increase is in the ratio of 1:4. This explains the extremely great difference in the chemical behaviour of helium and lithium. Lithium is thus monovalent and enters readily into chemical combination because its 2s electron is loosely bound to the nucleus and can be removed by an ionisation energy of only 5.39 eV while helium has the greatest ionisation potential (24.58eV) of all the elements. Likewise, all alkali elements (Na, K, Rb, Ce) have one electron in their outermost shell and hence are monovalent having similar chemical properties like lithium.

(4) Beryllium (Z = 4)

Beryllium, which comes after lithium in the second period of the table, has two electrons in the completed K-shell (n = 1). It has two additional electrons in the n = 2, ℓ = 0 subshell and is represented by $1s^22s^2$. Beryllium is one of the alkaline earth elements with a valence of 2. The atoms of the other elements of this group (Mg, Ca, Sr, Ba, Ra) should have similar structures. They have two electrons outside an inert gas or closed shell configuration and have very similar chemical and physical properties.

The electronic configurations of other elements in this period from boron (Z = 5) to neon (Z = 10) are

(i) boron (Z = 5). $1s^22s^22p$

(ii) carbon (Z = 6). $1s^22s^22p^2$

(iii) nitrogen (Z = 7). $1s^22s^22p^3$

(iv) oxygen (Z = 8). $1s^22s^22p^4$

(v) fluorine (Z = 9). $1s^22s^22p^5$

(vi) neon (Z = 10). $1s^22s^22p^6$

The last element of the second period, neon (Z = 10) have both its K and L - shells filled. Its two electrons fill up the K-shell and 8 electrons fill up the L - shell. Neon is therefore an inert gas. The radii of the atoms from lithium to neon, in which the L - shell (n = 2) is progressively filled steadily decreases as Z increases due to the

coulomb attraction between the additional electrons and the nucleus, while ionisation potential increases in the same order.

(5) Sodium (Z = 11)

$1s^2 2s^2 2p^6$ 3s. Sodium which lies in the third period has just one electron more than neon. This electron goes into 3s ($n = 3$, $\ell = 0$) sub-shell. It is an alkali metal like lithium, with a single valence electron. Because of the transition from quantum orbit $n = 2$ to $n = 3$, there is a marked increase in the radius of the atom and the ionization potential is much smaller. The electron can therefore be easily removed.

The next element in this period is magnesium (Z = 12). The 3s sub-shell of this atom is also filled in addition to K and L shells. These two electrons in the outermost incomplete M-shell make Mg a *divalent* element.

Argon (Z = 18) is the last element in this period. Its ten electrons fill up the K and L-shells, two electrons fill up the 3s sub-shell, six fill up the 3p sub-shell while the 3d sub-shell is vacant. It is an inert gas like helium and neon in the first and second period respectively. The electronic configuration of argon is $1s^2\, 2s^2\, 2p^6\, 3s^2\, 3p^6$.

Potassium (Z = 19), the first element in the fourth period, has just one electron more than argon. This electron does not go to 3d sub-shell, as expected, but goes to 4s sub-shell. This has been proved by the fact that the spectrum of potassium is similar to that of sodium with an s state as the ground state. Further the energy of potassium with the electron in 4s shell is less than that with this electron in the 3d sub-shell. Potassium has argon like core. The 3d sub-shell remains vacant until the element scandium (Z = 21).

The last element in the fourth period, Krypton (Z = 36) is as usual an inert gas. Its electrons fill up the K, L, M-shells and 4s and 4p sub-shells of N. Rubidium (Z = 37), the next element in the table, shows a disparity similar to that exhibited by potassium, namely that its 37^{th} electron, instead of going to the 4d and 4f sub-shells, goes to the 5s sub-shell. The 4d and 4f sub-shells remain vacant.

The distribution of electrons in atoms of many elements is given in Table. Let us consider the group formed by halogens, namely fluorine (Z = 9), chlorine (Z = 17), bromine (Z = 35) and iodine (Z = 53) (not included in the table). The outermost shells of each of these gases is incomplete by one electron. Each one of these

n, l		1, 0	2,0	2,1	3,0	3,1	3,2	4,0	4,1	4,2	4,3	5,0	5,1	5,2	5,3	5,4
X-ray symbols		K	L		M			N				O				
Electron symbols		1s	2r	2p	3s	3p	3d	4s	4p	4d	4f	5s	5p	5d	5f	5g
Element	Z															
H	1	1														
He	2	2														
Li	3	2	1	0												
Be	4	2	2	0												
B	5	2	2	1												
C	6	2	2	2												
N	7	2	2	3												
O	8	2	2	4												
F	9	2	2	5												
Ne	10	2	2	6												
Na	11	2	2	6	1	0										
Mg	12				2	0										
Al	13		Core of		2	1										
Si	14		10 electrons		2	2										
P	15		as in neon		2	3										
S	16	$1s^2$	$2s^2$	$3s^2$	2	4										
Cl	17				2	5										
A	18	2	2	6	2	6										

K	19	2	2	6	2	6	0	1	0				
Ca	20						0	2	0				
Sc	21						1	2	0				
Ti	22						2	2	0				
V	23						3	2	0				
Cr	24						5	1	0				
Mn	25	Core of 18 electrons as in argon $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$					5	2	0				
Fe	26						6	2	0				
Co	27						7	2	0				
Ni	28						8	2	0				
Cu	29						10	1	0				
Zn	30						10	2	0				
Ga	31						10	2	0				
Ge	32						10	2	1				
As	33						10	2	2				
Se	34						10	2	4				
Br	35						10	2	5				
Kr	36						10	2	6				
Rb	37	2	2	6	2	6	10	2	6	0	0	1	
Sr	38									0	0	2	
Y	39									1	0	2	
Zr	40									2	0	2	
Nb	41									4	0	1	
Mo	42	Core of 36 electrons as in krypton $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$ $3d^{10}$ $4s^2$ $4p^6$								5	0	1	
Te	43									6	0	1	
Ru	44									7	0	1	
Rh	45									8	0	1	
Pd	46									10	0	0	
Ag	47									10	0	1	
Cd	48									10	0	2	
In	49									10	0	2	1
Sn	50									10	0	2	2

gases can have the very stable structure of inert gases if they can have one more electron. Their tendency is therefore to acquire one more electron from somewhere. The halogens, therefore, readily enter into chemical combinations to make up for this deficiency. Elements like halogens, which tend to acquire an electron and become a negative ion, are known as *electro-negative.* Lithium, sodium potassium have only one valence electron and magnesium has two electrons in the outermost shell. These electrons are readily removed from the atom in chemical combinations, in thermionic or photo electric emission or in ionisation. These and such other elements are called *electro-positive*, because they tend to get rid of one or more electrons and become positive ions. Electro-positive elements have smaller number of electrons in the outermost shell or sub-shell than the maximum number of electrons possible in them, while the electro-negative elements have almost but not completely filled outermost shells.

EXERCISES

1. Describe the experimental method for investigating scattering of α-particles by matter and derive Rutherford's formula for it.

2. Give an account of Rutherford's theory of scattering of α-particles and describe how it has been verified.

3. Describe Rutherford Atom model and the evidence that led to it. What are the drawbacks of Rutherford atom model?

4. State the postulates of Bohr atom model. Obtain expressions for the radius and electron energy in the n^{th} orbit.

5. Write down the postulates of Bohr atom model for hydrogen atom. In what way do they violate the classical laws of electromagnetic theory?

6. Give an account of Bohr's theory of hydrogen spectrum taking into account the finite mass of hydrogen nucleus. Derive the formula for the Balmer series of hydrogen atom.

7. Derive an expression for the energy levels of hydrogen atom in Bohr model clearly stating the physical significance of the basic step.

   Find the expression for the Rydberg constant. Show how the limits of Lyman and Balmer series are related to each other.

8. Define the terms (i) critical potential, (ii) excitation potential and (iii) ionisation potential. Describe Frank-Hertz experiment for determining the critical potentials. Show how the results of the experiment confirm Bohr's postulates.

9. Give an account of Sommerfeld's extension of Bohr atom model. How does it account for the fine structure of hydrogen atom?

10. Find an expression for the velocity of electron in the n^{th} Bohr orbit.

11. Discuss the general characteristics of Sommerfeld's elliptical orbits and show that the total energy of an electron moving in Sommerfeld's orbits of the same total quantum number is the same and is identical with that of corresponding to Bohr orbits. How is the result modified by introducing correction for relativistic variation of electron mass?

12. Find the frequency of revolution of an electron in the first Bohr orbit of the hydrogen atom.

13. Describe vector model of the atom and explain the different quantum numbers associated with it. Give two important applications of the model.

14. Discuss the salient features of the vector model of the atom. Write down the electronic configuration of Cu (Z = 29).

15. Explain the physical significance of the different quantum numbers used in the vector model of the atom.

16. State and explain Pauli's exclusion principle as applied to electrons in atoms. Describe how this principle assists in the interpretation of the periodic system of the elements.

17. What is Pauli's exclusion principle? Explain the configuration of electrons on the basis of this principle. Give the electronic configuration of cobalt (Z = 27).

18. Write short notes on.
    (i) Thomson atom model
    (ii) Correspondence principle
    (iii) L-S coupling and J-J coupling
    (iv) Vector atom model
    (v) Energy level diagram of hydrogen atom
    (vi) Pauli's exclusion principle and periodic table of elements
    (vii) Rydberg-Ritz combination principle.

19. What is the impact parameter of a 5 MeV alpha particle scattered through 10° when it approaches a gold nucleus (Z = 79)? **[2.6 x 10^{-11} cm]**

20. Determine the distance of closest approach of 1 MeV protons incident upon gold nuclei. **[1.14 x 10^{-11} cm]**

21. A 5 MeV alpha-particle approaches a gold nucleus with an impact parameter of 2.6 x 10^{-13}m. Through what angle will it be scattered? **[10°]**

22. What fraction of a beam of 7.7 MeV alpha-particles incident upon a gold foil is scattered by less than 1°? **[0.876]**

23. For equal number of incident particles, determine the ratio of protons to alpha-particles of the same energy that will be scattered by angles greater than 90° from a gold foil. **[0.25]**

24. Taking the wavelength of H_α to be 6563A, calculate the Rydberg constant for hydrogen. **[1.097 x 10^7 m^{-1}]**

25. Why the quantum number n = 0 was not allowed in Bohr atom model? [If n = 0, then r = 0, *i.e.*, the electron would have to go through the nucleus, which is impossible]

26. If the series limit of Balmer series lies at 3646 A, calculate the wavelength of the first members of lyman and Paschen series. **[1215 A, 18760 A]**

27. Find the wavelength of the photon emitted when the hydrogen atom goes from n = 10 state to the ground state. **[921 A]**

28. The Rydberg constant for hydrogen is 10967700 m^{-1}. Calculate the short and long wavelength limits of Lyman series. **[911.6 A, 1215 A]**

29. Calculate the radii of the first, second and third permitted electron orbits in a hydrogen atom. **[0.527 A, 2.108 A, 4.743 A]**

30. What are the H_α, H_β, H_γ and H_δ lines in the spectrum of hydrogen atom? Explain their origin. Derive an expression for Rydberg constant and calculate its numerical value using the constants given below:

    e = 1.6 x 10^{-19} coulomb

    h = 6.625 x 10^{-34} J-s

    m = 9.1 x 10^{-31} Kg

31. If the Rydberg constant is $1.097 \times 10^7\ m^{-1}$, what are the wavelengths of the first three lines of the Paschen series? **[18750A, 12810A, 10930A]**

32. Calculate the radius of the second Bohr's orbit of the hydrogen atom and hence find out the total energy of the electron in that orbit. ($h = 6.6256 \times 10^{-34}$ J-s, $m = 9.1 \times 10^{-31}$ Kg, $e = 1.6 \times 10^{-19}$ coulomb). **[2.117×10^{-10} m, 3.4 eV]**

33. Show that the second ionization potential of helium is 54.34 volts, given $e = 1.6 \times 10^{-19}$ coulomb, $m = 9.11 \times 10^{-31}$ Kg, $h = 6.62 \times 10^{-34}$ J-s.

34. A photon of wavelength 1400 A is absorbed by cold mercury vapour and two other photons are emitted. If one of them is of wavelength 1850A, what is the wavelength of the other line?

    [Energy absorbed, $E = \frac{hc}{\lambda} = \frac{(6.62\times10^{-34})(3\times10^{8})}{(1400\times10^{-10})} J = 8.867$ eV.

    1850 A line absorbs 6.71 eV. The balance (8.867 – 6.71) = 2.157 eV corresponds to a line of wavelength 5755A]

35. The wavelength of the second line of the Balmer series in the hydrogen spectrum is 4861A. Calculate the wavelength of the first linc. [6563A]

36. The ionisation potential of an atom is 14.2eV. Calculate the series limit in its absorption spectrum. [$\lambda = \frac{hc}{eV} = 871.5$A]

37. In Frank-Hertz experiment, the excitation and ionisation potentials for sodium are 2.1, 3.7 and 5.13 volts respectively. Calculate the velocities of the impinging electrons to excite and ionize the sodium atoms. [$\frac{1}{2}mv^2 = eV$; $v = \sqrt{\frac{2eV}{m}}$; $v_1 = 8.5592 \times 10^5\ ms^{-1}$; $v_2 = 11.40 \times 10^5\ ms^{-1}$; $v_3 = 13.43 \times 10^5\ ms^{-1}$]

38. Calculate the critical potential in volts corresponding to the excitation of mercury line $\lambda = 5416$A. [$V = \frac{hc}{e\lambda}$; 2.29 volts]

39. Ultraviolet light of 800A is incident on hydrogen in a quartz tube. Calculate the kinetic energies with which electrons will be ejected from the hydrogen atoms. Express the results in electron volts. **[1.85eV; 12.06 eV; 13.94 eV; 14.60 eV]**

CHAPTER XV

WAVE MECHANICS

Introduction

The phenomena of interference, diffraction and polarization of light could be explained on the basis of wave nature of light. However, there are certain phenomena like photo-electric effect. Compton effect, etc, which could not be explained by the wave theory of light–these could be explained only by the photon theory of light, *i.e*, corpuscular nature of radiation. Thus a dual nature, that of wave and particle, came to be associated with light. In 1924, Louis de Broglie put forward the suggestion that matter, like radiation, has dual nature, *i.e*, matter which is made of discrete particles (atoms, protons, electrons, etc) might exhibit wavelike properties under appropriate conditions.

15.1 de Broglie's concept of matter waves

The considerations that led de Broglie to develop his concept of matter waves are as follows:

(i) Nature loves symmetry. Matter and energy, the two fundamental forms in which nature manifests itself, must be mutually symmetrical. Radiant energy possesses dual nature – wave and particle; matter must also possess the same dual nature – particle and wave.

(ii) Close parallelism between mechanics and optics as regards their fundamental laws governing motion, momentum, mass and energy.

Maupertian principle of least action states that a moving particle always chooses that path for which action is minimum *i.e*, the integral of the momentum over the path is a minimum and is analytically expressed as

$$\partial \int_{P_1}^{P_2} (mu)ds = 0$$

Similarly, **Fermat's principle of least time in optics implies similar** condition. The principle states that a light ray always chooses that path for which the time of transit is a minimum and is analytically expressed as

$$\partial \int_{P_1}^{P_2} \mu ds = 0$$

The close analogy of the two principle belonging to two different branches of physics suggest the probability that matter, under suitable condition, might behave as wavelike entity.

(iii) Bohr's theory of atomic structure. According to Bohr's theory of atomic structure, the radii of the privileged orbits are proportional to the square of integral numbers [$r = (\epsilon_0 h^2 n^2)/\pi m_e Z e^2$]. Thus the stable non-radiating orbits of the electron in the atom are governed by *integer rules.* The only other phenomena where integer rules are applied are *interference* and modes of *vibrations of stretched strings* – both of which are wave phenomena. Hence, electrons in the privileged orbits could not be regarded simply as material particles but a certain intrinsic periodicity (*i.e*, wavelike nature) should also be assigned to them.

From these considerations and from his own intuitive feeling, Louis de Broglie made, in 1924, a very bold and novel suggestion that like radiation, matter also has dual (*i.e*, particle–like and wave-like) characteristics. According to this hypothesis, moving particle of matter like electron, proton, neutron, atoms or molecules, has always got a wave associated with it and the particle is controlled by the wave in a manner similar to that in which a photon is controlled by waves. While considering its mass, charge and energy, the electron is considered as a particle, but when the path of a beam of electron or the reflection of a beam of electron is considered, electron is treated as though it were a beam of waves. *These waves, associated with a material particle, are called matter waves.* For obvious reasons they are also known as *de Broglie waves.*

The de Broglie wavelength

The wavelength associated with a particle, say an electron, is referred to as de Broglie wavelength. This may be computed as follows:

The particle may be pictured as a standing wave system in the region of space occupied by the particle. Let the quantity that undergoes periodic changes giving rise to matter waves be ψ. Let its value at a point x_0, y_0, z_0 at any instant of time t_0 in the immediate vicinity of the particle be given by

$$\psi = \psi_0 \sin 2\pi\nu_0 t_0$$

where ψ_0 is the amplitude at the point chosen and ν_0 the frequency as observed by an observer at rest with respect to the particle.

Let the particle be given a velocity v along the x-axis. In order to represent the variation of ψ under this new condition, we have to apply the *transformation equation* of relativity

$$t_0 = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - v^2/c^2}}$$

$$\therefore \quad \psi = \psi_0 \sin 2\pi\nu_0 \frac{t - \frac{vx}{c^2}}{\sqrt{1 - v^2/c^2}}$$

This equation may be compared with

$$y = A \sin \frac{2\pi}{T}\left(t - \frac{x}{u}\right)$$

where A is the amplitude, T the time period and u the velocity of the wave along the x-direction. Comparing, we get

$$u = \frac{c^2}{v} \text{ and } \frac{1}{T} = \nu = \frac{\nu_0}{\sqrt{1 - v^2/c^2}}$$

From Einstein's mass-energy relation

$$h\nu_0 = m_0 c^2 \text{ ; or } \nu_0 = \frac{m_0 c^2}{h}$$

$$\therefore \nu = \frac{m_0 c^2}{h\sqrt{1 - v^2/c^2}}$$

$$\text{But } \frac{m_0}{\sqrt{1 - v^2/c^2}} = m; \quad \therefore \nu = \frac{mc^2}{h}$$

Hence the wavelength of the matter wave,

$$\lambda = \frac{\text{velocity}}{\text{frequency}} = \frac{u}{\nu} = \frac{c^2/v}{mc^2/h} = \frac{h}{mv} \qquad (15.1)$$

The above relation is known as *de Brglie wave equation* and the wave associated with the particle as *de Broglie wavelength.*

de Brglie wavelength of an electron

Let an electron be accelerated through a potential difference of V volts. Then the kinetic energy of the electron is

$$\frac{1}{2}mv^2 = eV$$

where e is the electronic charge.

$$\text{or} \quad m^2v^2 = 2meV$$

$$\text{or} \quad mv = \sqrt{2meV}$$

Hence the wavelength associated with the wave is

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2meV}}$$

Ignoring relativistic corrections, $m = m_0$

$$\therefore \quad \lambda = \frac{h}{\sqrt{2em_0V}}$$

Wave number

The wave number ($\overline{\nu}$) of a wave is given by the reciprocal of its wavelength.

$$\text{or,} \quad \overline{\nu} = \frac{\sqrt{2em_0V}}{h} m^{-1}$$

Characteristics of the de Broglie matter waves

(i) Since $\lambda = h/\sqrt{2emV}$, it is seen that the larger the mass of the matter particle, the shorter is the associated wavelength.

(ii) These waves are not electromagnetic waves but are a new kind of wave. They may be regarded as *pilot* waves in the sense that their only function is to pilot or guide the matter particle. The component waves of de Broglie wave packet have amplitudes which vary with the likelihood of detecting the particle at a particular position as depicted in the schematic diagram of the wave packet shown in Fig. 15.1.

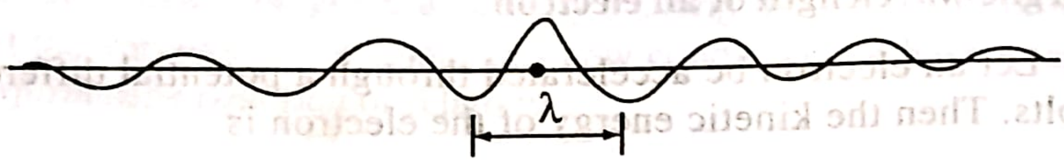


Fig. 15.1

(iii) Two different velocities are associated with a material particle in motion – one refers to the mechanical motion of the particle (v) and the other refers to the propagation of the associated wave (u).

The phase or the wave velocity of the matter wave is

$u = \nu\lambda$ where ν is the frequency.

By photon analogy

$$E = h\nu; \qquad \text{or} \quad \nu = \frac{E}{h}$$

$$\text{Also } E = mc^2 \qquad \therefore \quad \nu = \frac{mc^2}{h}$$

$$\text{Now } \lambda = \frac{h}{mv} \qquad \therefore \quad u = \nu\lambda = \frac{mc^2}{h} \cdot \frac{h}{mv} = \frac{c^2}{v}$$

Since the speed v of the material particle cannot exceed the velocity of light, it is obvious that the phase velocity of the wave associated with a material particle in motion is greater than the velocity of light.

15.2 Wave velocity and Group velocity for de Broglie waves

According to de Broglie's concept of matter waves, each particle of matter (like electron, proton, etc), while in motion, may be regarded as consisting of a ***group of waves*** or a ***wave packet*** as it is called. The wave packet, formed by the superposition of a number of waves and traveling with the velocity of the particle, behaves very much like a corpuscle. Each component wave propagates with a definite velocity,

called the *wave velocity* or *phase velocity*. But when a disturbance consists of a number of component waves, each traveling with slightly different velocity, the resultant velocity will be that of a *periodicity*. This velocity with which the periodicity advances is called the *group velocity*. It can be shown that this wave group associated with a moving particle travels with the same velocity as the particle.

To show that the particle velocity (v) is the same as the group velocity (w).

The group velocity (w) of a system of waves is given by

$$w = u - \lambda \frac{du}{d\lambda} \tag{15.2}$$

where λ is the wavelength of an individual wave in the system and u the velocity of that wave, frequently called the *phase velocity* or *wave velocity*. In a dispersive media where $du/d\lambda$ is positive, $w < u$.

In free space $du/d\lambda = 0$ and $w = u$.

Eqn. (15.2) may be written as

$$w = \lambda^2\left(\frac{u}{\lambda^2} - \frac{1}{\lambda}\frac{du}{d\lambda}\right)$$

$$= -\lambda^2 \frac{d}{d\lambda}\left(\frac{u}{\lambda}\right) = -\lambda^2 \frac{d\nu}{d\lambda}$$

where ν is the frequency of the wave.

$$\therefore \quad \frac{1}{w} = -\frac{1}{\lambda^2}\frac{d\lambda}{d\nu} = \frac{d}{d\nu}\left(\frac{1}{\lambda}\right) \tag{15.3}$$

Let E = total energy of the particle and V = potential energy of the particle.

Then the kinetic energy of the particle

$$= \frac{1}{2}mv^2 = E - V$$

$$\text{or} \quad v = \{2(E - V)/m\}^{\frac{1}{2}} \tag{15.4}$$

For matter waves, $\lambda = h/mv$

$$\therefore \quad \frac{1}{\lambda} = \frac{mv}{h} = \frac{1}{h}\{2(E-V)/m\}^{\frac{1}{2}} \qquad (15.5)$$

Substituting this value of 1/λ in (15.3),

$$\frac{1}{w} = \frac{d}{dv}\{\frac{1}{h}\sqrt{2m(E-V)}\}$$

$$= \frac{d}{dv}\{\frac{1}{h}\sqrt{2m(hv-V)}\}$$

$$\frac{1}{w} = \frac{1}{h}\frac{d}{dv}\{2m(hv-V)\}^{\frac{1}{2}}$$

$$= \frac{1}{h} \times \frac{1}{2}\{2m(hv-V)\}^{-\frac{1}{2}} \times 2mh$$

$$\therefore \quad \frac{1}{w} = \frac{m}{\{2m(hv-V)\}^{\frac{1}{2}}}$$

$$= \left\{\frac{m}{2(E-V)}\right\}^{\frac{1}{2}} = \frac{1}{v} \qquad (15.6)$$

∴ **the group velocity (w) = the particle velocity (v).**

***Example 15.1** Compute the de Broglie wavelength of the following:*

(i) A 1000-kg automobile traveling at 100 m/s (about 200 miles/hour).

(ii) A 10-gm bullet traveling at 500m/s.

(iii) A smoke particle of mass 10^{-9} gm moving at 1 cm/s.

(iv) An electron with a kinetic energy of 1eV.

(v) An electron with a kinetic energy of 100 MeV.

Soln.

(i) $$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{6.6\times10^{-34}\,J.s}{(1000kg)(100m/s)} = 6.9 \times 10^{-39}\,m$$

$$= 6.9 \times 10^{-29}\,\mathring{A}$$

(ii) $$\lambda = \frac{h}{mv} = \frac{6.6\times10^{-34}\,J.s}{(10\times10^{-3}\,kg)(500m/s)} = 1.3 \times 10^{-34}\,m$$

$$= 1.3 \times 10^{-24}\,\mathring{A}$$

(iii) $$\lambda = \frac{h}{mv} = \frac{6.6\times10^{-34}\,J.s}{(10^{-9}\times10^{-3}\,kg)(1\times10^{-2}\,m/s))} = 6.6 \times 10^{-20}\,m$$

$$= 6.6 \times 10^{-10}\,\mathring{A}$$

(iv) The rest mass energy (m_0c^2) of an electron is 5.11×10^5eV. Since the kinetic energy of 1eV is much less than the rest mass energy, we can use non-relativistic kinematics.

$$p = mv = \sqrt{2meV} = \sqrt{2mK}$$

where K = kinetic energy = eV.

$$\therefore \quad mv = \sqrt{2mK} = \sqrt{(2)(9.1\times10^{-31}\,kg)(1eV)(1.6\times10^{-19}\,J/eV)}$$

$$= 5.4 \times 10^{-25}\ kg.m/s.$$

Then $$\lambda = \frac{h}{mv} = \frac{6.6\times10^{-34}\,J.s}{5.4\times10^{-25}\,kgm/s} = 1.2\times10^{-9}\,m = 1.2nm = 12\mathring{A}$$

The above solution can also be obtained in the following way:

$$p = \sqrt{2mK} = \sqrt{\frac{2(mc^2)K}{c^2}} = \frac{1}{c}\sqrt{2(mc^2)K}$$

or $\quad cp = \sqrt{(2)(5.1\times10^5 eV)(1eV)} = 1.0\times10^3 eV$

Then $\lambda = \frac{h}{p} = \frac{hc}{pc}$

Now $hc = (6.6\times10^{-34}\ J.s)(3\times10^8\ m/s)$

$= (6.6\times3\times10^{-26})\ J.m$

$= \frac{6.6\times3\times10^{-26}}{1.6\times10^{-19}} eV.m$

$= 12.4 \times 10^{-7}\ eV.m$

$= 1240 \times 10^{-9}\ eV.m = 1240\ eV.nm$

$\therefore \quad \lambda = \frac{hc}{pc} = \frac{1240 eV.nm}{1.0\times10^3 eV} = 1.2nm$

(v) In this case, the kinetic energy is much greater than the rest mass energy. So the total energy $E \cong$ kinetic energy K.

In this case we can use the extreme relativistic approximation

$$E = \sqrt{(pc)^2 + (mc^2)^2} \cong pc$$

$\therefore \quad \lambda = \frac{hc}{pc} = \frac{1240 eV.nm}{100 MeV}$

$= \frac{1240 MeV.fm}{100 MeV} = 12.4 fm$

***Example 15.2** Calculate the de Broglie wavelength of a proton moving with $1/20^{th}$ of the velocity of light.*

Soln.

Velocity of proton, $v = \frac{c}{20} = \frac{3\times10^8}{20} ms^{-1}$

mass of proton, $m = 1.673 \times 10^{-27} kg$

$$\lambda = \frac{h}{mv} = \frac{(6.624\times10^{-34}\,J.s)(20)}{(1.673\times10^{-27})(3\times10^{8})}$$

$$= 2.640 \times 10^{-14}\ m.$$

Example 15.3 Calculate the wavelength associated with a thermal neutron of energy 0.025 eV.

Soln.

Mass of the neutron, $m = 1.676 \times 10^{-27}$ kg

$$\text{energy of the neutron} = \frac{1}{2}mv^2 = (0.025\times1.6\times10^{-19}\,J)$$

$$1eV = 1.6\times10^{-19}\,J$$

$$\therefore\quad mv = [\,2\times 1.676 \times 10^{-27} \times 0.025\times1.6\times10^{-19}]^{1/2}$$

$$\therefore\quad \lambda = \frac{h}{mv} = \frac{6.62\times10^{-34}\,J.s}{[2\times1.676\times10^{-27}\times0.025\times1.6\times10^{-19}]^{1/2}}$$

$$= 1.807 \times 10^{-10}\ m.$$

Example 15.4 An electron initially at rest is accelerated through a p.d. of 5000V. Compute (i) the momentum, (ii) the de Broglie wavelength and (iii) the wave number of the electron wave.

Soln.

(i) momentum of the electron is

$$p = mv = \sqrt{2emV} = [2\times1.60\times10^{-19}\times9.1\times10^{-31}\times5000]^{1/2}$$

$$= 3.82 \times 10^{-23}\ kg.m/s$$

(ii) $$\lambda = \frac{12.25}{\sqrt{5000}}\overset{0}{A} = 0.175\overset{0}{A}$$

(iii) $$\text{wave number, } \nu = \frac{1}{\lambda} = \frac{1}{0.175\times10^{-10}} = 5.71\times10^{10}.$$

Example 15.5 Compute the de Broglie wavelength of a proton whose kinetic energy is equal to the rest energy of an electron. Mass of a photon is 1836 times that of the electron.

Soln.

According to Einstein's mass energy relation, the rest energy of an electron is

$$m_0c^2 = (9.1\times10^{-31})(3\times10^8)^2 = 81.9\times10^{-15}\,\text{J}.$$

$$\text{mass of proton} = 1836\times 9.1\times 10^{-31}\ \text{kg}$$

$$= 1.671\times 10^{-27}\ \text{kg}$$

$$\frac{1}{2}mv^2 = 81.9\times10^{-15}$$

$$v = \left[\frac{(2)(81.9\times10^{-15})}{1.671\times10^{-27}}\right]^{1/2} = 9.9\times10^6\,\text{m}$$

$$\therefore\quad \lambda = \frac{h}{mv} = \frac{6.62\times10^{-34}}{(1.671\times10^{-27})(9.9\times10^6)}$$

$$= 4\times10^{-14}\,\text{m}$$

$$= 0.0004\,\text{Å}$$

15.3 Experimental study of matter waves

The material particles, which readily lend themselves to experimental investigation of matter waves, are the electrons. Not only are they easily produced in fairly intense beams with a definite velocity, like the cathode rays, but because of the smallness of their mass, the wavelengths associated with them might have magnitudes which could be measurable. The de Broglie wavelength of the electron is given by

$$\lambda = \frac{h}{\sqrt{2emV}}$$

$$= \frac{6.625\times10^{-34}}{[2\times(1.6\times10^{-19})(9.11\times10^{-31})V]^{1/2}}$$

$$= \frac{12.27}{\sqrt{V}}\times10^{-10}\,m$$

$$= \frac{12.27}{\sqrt{V}}\,\mathring{A}$$

Thus the wavelength of the de Broglie waves associated with electrons accelerated through a P.D. of 100V in vacuum is $1.227\,\mathring{A}$. This is comparable with the wavelength of X-rays: the de Broglie waves should be capable of being diffracted by a crystal, just as X-rays are.

Experiments of Davisson and Germer

The first experimental confirmation of the wave nature of electrons (and the quantitative confirmation of the de Broglie relationship $\lambda = h/p$) followed soon after de Broglie's original hypothesis. In 1926, at the Bell Telephone Laboratories, Clinton Davisson and Lester Germer were investigating the reflection of electron beams from the surface of nickel target. As frequently happens an experimental accident was to result in a new discovery the nickel target was accidentally subjected to such heat treatment that it was transformed into a group of large crystals. As a result, the reflection of electrons became anomalous, *i.e*, instead of decreasing continuously from the angular position of maximum reflection, the reflected intensity showed striking maxima and minima. This unexpected result called to mind X-ray diffraction from crystals and made them suspect that a beam of electrons might be diffracted from crystal like X-rays. This would mean that under certain conditions, electrons behave like waves. This was a very important point which deserved verification. So they prepared a target consisting of a single crystal of nickel and carried out the following experiment, which established the fact that a beam of electrons could be reflected, diffracted or even refracted.

Experimental arrangement

The experimental arrangement is shown diagrammatically in Fig.15.2. Electrons are produced from what is known as an **electron gun G.** This contains a tungsten filament F heated to dull red, when electrons are emitted by thermionic action. The electrons are then accelerated by passing them through an electric field of known potential difference and collimated by suitable slits so that a fine parallel beam of electrons emerges. The electron beam is then directed to fall on a large single crystal of nickel, known as the target T which is capable of rotation about an axis parallel to the axis of the incident beam. The electrons are *reflected* from the crystal in different directions, the angular distribution being measured by a Faraday cylinder, called the *collector C*. The collector is connected to a sensitive

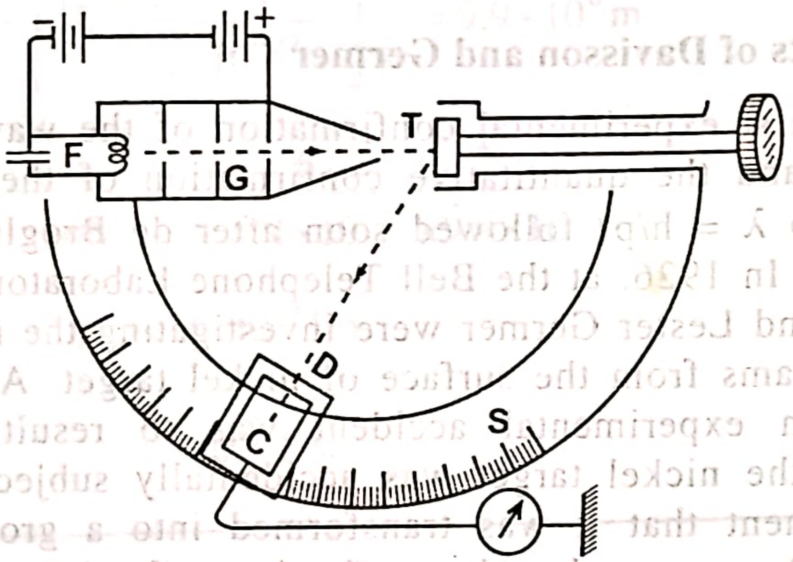


Fig. 15.2

galvanometer and can be moved along a sensitive graduated circular scale S, so that it is able to receive the reflected electrons and measure the intensity of the electron beam scattered at all angles between 20° and 90°. The collector has two walls insulated from each other. A retarding potential is applied between the inner wall c and the outer wall D so that only the electrons possessing nearly the incident velocity and not the secondary slow electrons, excited by collisions with atoms, may enter the collector and be detected by the galvanometer. The accelerating potential used ranged from about 30 to 600 volts and the retarding potential was *nine-tenths* of the accelerating voltage. The whole apparatus was completely enclosed, highly evacuated and degassed.

The nickel crystal which belongs to the face centred cubic type, is so cut as to present a smooth reflecting surface parallel to the lattice plan (111), *i.e.*, perpendicular to one of the diagonals of the cube. By rotating the crystal, any azymuth of the crystal can be presented to the plane defined by the incident beam and the reflected beam entering the collector.

Experimental procedure: The experiment was conducted in the following two different ways:

(i) *Normal incidence.* Here the beam of the electrons is allowed to fall normally on the surface of the crystal, which acts as a plane grating and produces a diffraction effect. For each azymuth of the crystal, a beam of electrons is made to fall normally on the crystal. The collector was moved to various positions on the scale S and the galvanometer deflection was noted at each position. The deflection was a measure of the intensity of the diffracted beam of electrons. The galvanometer deflection is plotted against the angle between the incident beam and the beam entering the collector, known as the *colatitude.* The observations were repeated for different accelerating voltages and a number of curves were drawn as shown in Fig. 15.3. As can be seen in the figure, for an accelerating potential of 40 volts, a smooth curve was obtained. At 44 volts, a small but distinct bump appears at an angle of about 60° (Fig. 15.3a).The bump increases at higher accelerating voltage and the angle at which it appears also changes. The bump becomes most prominent at an accelerating voltage of 54 volts and appears at angle of 50°. At still higher voltages, the bump gradually diminishes, there being hardly any trace of it at about 68 volts.

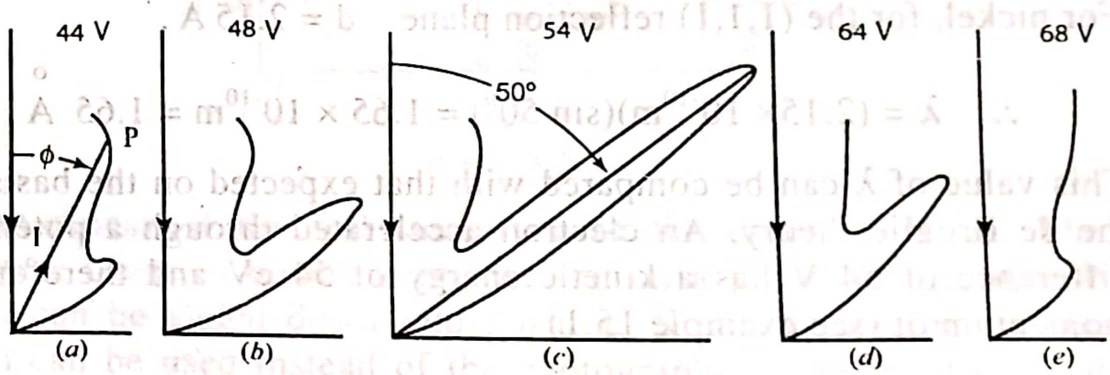


Fig. 15.3

Because the electrons were of low energy, they did not penetrate very far into the crystal, and it is sufficient to consider the diffraction to take place in the plane of atoms on the surface. The situation is entirely similar to using a reflection grating for light, the spacing d between the rows of atoms on the crystal is analogous to the spacing between the slits in the optical grating (Fig.15.4). The maxima for a diffraction grating occur at angles ϕ such that

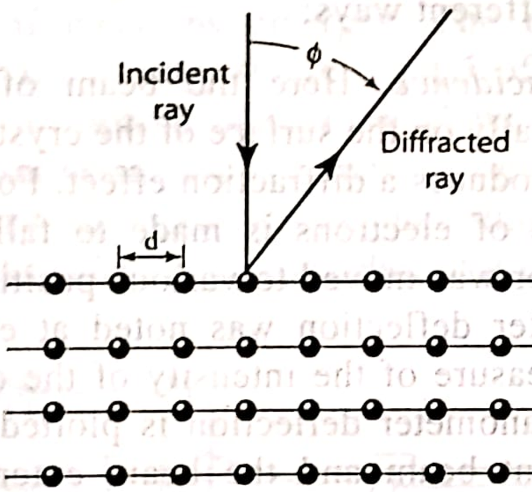


Fig. 15.4

$$d \sin \phi = n\lambda \tag{15.7}$$

where n (= 1, 2, 3, ...) is the order of the maximum. The peak at ϕ = 50° must be a first-order peak (n = 1), because no peaks were observed at smaller angles. If this is indeed an interference maximum, the corresponding wavelength is, from eqn. 15.7

$$\lambda = d \sin \phi$$

For nickel, for the (1,1,1) reflection plane $\quad d = 2.15\ \mathring{A}$.

$$\therefore \quad \lambda = (2.15\times 10^{-10}\text{m})(\sin 50^0) = 1.65 \times 10^{-10}\text{m} = 1.65\ \mathring{A}$$

This value of λ can be compared with that expected on the basis of the de Broglie theory. An electron accelerated through a potential difference of 54 V has a kinetic energy of 54 eV and therefore a momentum of (see example 15.1)

$$p = \sqrt{2mk} = \frac{1}{c}\sqrt{2mc^2k} = \frac{1}{c}(7430\text{eV})$$

The de Broglie wavelength is given by

$$\lambda = \frac{h}{p} = \frac{hc}{pc}, \text{ using } hc = 1240 \text{ eV. nm}$$

$$\lambda = \frac{1240\text{eV.nm}}{7430\text{eV}} = 0.167\text{nm} = 1.67\,\overset{o}{\text{A}}$$

This is an excellent agreement with the value found from the diffraction maximum, and provides strong evidence in favour of the de Broglie theory. For this experimental work, Davisson shared the 1937 Nobel prize with G.P Thomson.

15.4 The experiments of G.P. Thomson

In 1928 G.P. Thomson in Scotland extended the research on electron waves to high speed electrons ranging from 10,000 to 50,000 volts, diffracted by very thin metallic films. His method is analogous to the Debye-Scherrer powder method of X-ray analysis of crystals.

Experimental arrangement. The experimental arrangement is shown in Fig.15.5. A beam of cathode rays is produced in a discharge tube AC by means of an induction coil. The rays are passed through a diaphragm tube A to obtain a fine pencil of electrons which is then allowed to fall upon a very thin gold foil F.

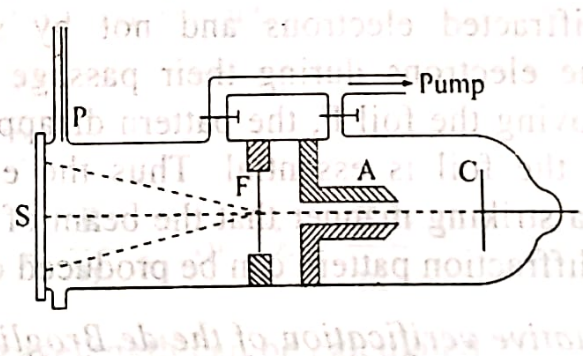


Fig. 15.5

The thickness of the foil should be of the order of 10^{-8}m. The emergent beam of electrons is received on a photographic plate P which can be slided down into position. S is a fluorescent screen which can be used instead of the photographic plate for the visual examination of the result obtained by the passage of electrons through the foil. The camera part FP of the apparatus is exhausted to

a very high vacuum while air is allowed to leak into the discharge tube section through a needle valve. Since the only connection between the camera and the discharge tube is through the diaphragm A, it is possible to have the camera at a low pressure and yet maintain the discharge tube sufficiently soft to give a beam of the required voltage.

Experimental procedure. A beam of electrons of known velocity is allowed to fall on the photographic plate after traversing the thin foil. When the plate is developed, a symmetrical pattern consisting of concentric rings about a central spot is obtained (Fig. 15.6) very much like that produced by X-rays in the powdered crystal method. When the cathod rays in the discharge tube are deflected by means of a magnetic field, it is found that the entire pattern on the fluorescent

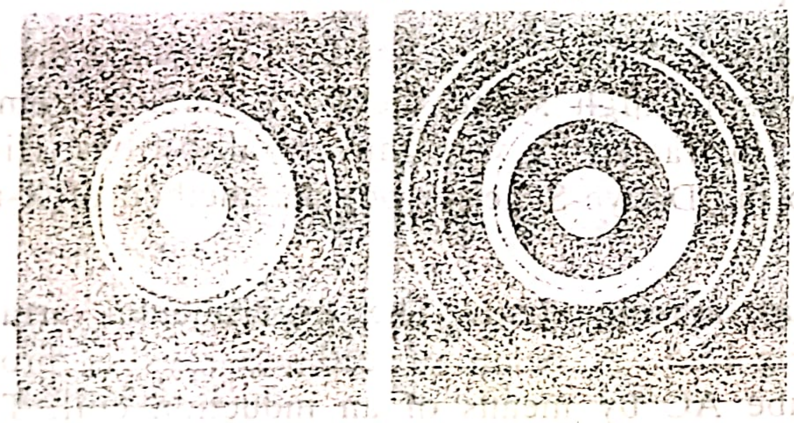


Fig. 15.6

screen S shifts correspondingly. This confirms that the pattern is produced by diffracted electrons and not by secondary X-rays generated by the electrons during their passage through the foil. Further, on removing the foil F, the pattern disappears showing that the presence of the foil is essential. Thus the experiment clearly demonstrates in a striking manner that the beam of electrons behaves as waves, since diffraction pattern can be produced only by waves.

The quantitative verification of the de Broglie equation can be made as follows:

As with the Debye-Scherrer powder crystal method for X-rays, in the polycrystalline film there will be some crystal set at the correct angle to give a Bragg reflection. If there are enough crystals distributed at random, the result of such reflections will be a series of rings, arising from the intersection of the cones of diffraction with the photographic plate. In Fig.15.7 let AB be the incident beam

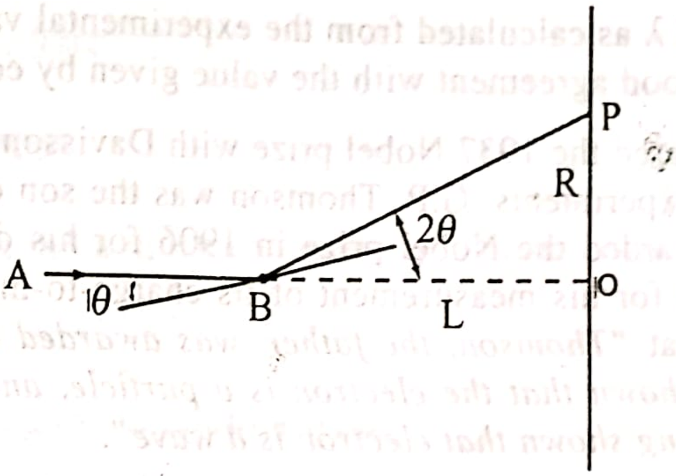


Fig. 15.7

passing through the film at B and let BP be the direction of the beam which has suffered a Bragg reflection in some crystals in the film at B. This reflected beam falls at the point P on the photographic plate at a distance R from the central point O. Let the distance BO from the film to the plate be L. The angle PBO = 2θ, where θ is given by the Bragg relation $n\lambda = 2d \sin\theta$.

$$\text{Now } \frac{R}{L} = \tan 2\theta = 2\theta, \text{ since } \theta \text{ is small.}$$

$$\therefore \quad \theta = \frac{R}{2L}$$

$$\text{But } 2d \sin\theta = 2d.\ \theta = n\lambda$$

$$\text{or} \quad \theta = \frac{n\lambda}{2d}$$

$$\therefore \quad \frac{n\lambda}{2d} = \frac{R}{2L}; \quad \text{or} \quad \lambda = \frac{Rd}{nL}$$

From this the wavelength can be calculated.

G.P. Thomson employed very high voltages of the order of 50,000 volts to accelerate the electrons. Applying relativistic correction for the mass for very high speed electron, it can be shown that

$$\lambda = \frac{12.27}{\sqrt{V}} \cdot \frac{1}{\sqrt{1+9.836\times10^{-7}\,V}}\ \mathring{A} \qquad (15.8)$$

The value of λ as calculated from the experimental values of R, D, L and n is in good agreement with the value given by eqn.15.8.

Thomson shared the 1937 Nobel prize with Davisson for his electron diffraction experiments. G.P. Thomson was the son of J.J.Thomson, who was awarded the Nobel prize in 1906 for his discovery of the electron and for his measurement of its charge-to-mass ratio. It has been said that "*Thomson, the father, was awarded the Nobel prize for having shown that the electron is a particle, and Thomson, the son, for having shown that electron is a wave*".

Example 15.6 *An electron, initially at rest, is accelerated through a P.D. of 5000V. Compute (i) the momentum and (ii) the de Broglie wavelength of the electron. (iii) Also calculate the Bragg angle for it's first order reflection from the (1 1 1) plane of nickel which are 2.04 Å apart (iv) What is the angle of deviation of the first order diffraction maximum?*

Soln.

Let v be the velocity of the electron. Then its K.E = $\frac{1}{2}mv^2 = V.e$

(i) Momentum of the electron,

$$mv = \sqrt{2mVe} = \sqrt{(2)(9.1\times10^{-31}kg)(5000V)(1.6\times10^{-19}C)}$$

$$= 3.818 \times 10^{-23}\ kg.ms^{-1}$$

(ii) $$\lambda = \frac{h}{mv} = \frac{6.625\times10^{-34}J.s}{3.818\times10^{-23}kg.ms^{-1}} = 0.1729\times10^{-10}m$$

$$= 0.1729\ \text{Å}.$$

(iii) Let θ be the Bragg angle for the first order reflection from (111) planes of nickel. Now by Bragg's relation,

$$2d\sin\theta = n\lambda.$$

Here n=1, $\lambda = 0.1729\times10^{-10}m$, d= $2.04 \times 10^{-10}m$

$\therefore \quad (2)(2.04 \times 10^{-10}) \sin\theta = (1)(0.1729 \times 10^{-10})$

or $\quad \sin\theta = 0.04237$

or $\quad \theta = 2°25'$

(iv) Angle through which the electron is deviated

$= 2\theta = (2)\ (2°25') = 4°50'.$

15.5 Wave Mechanical Atom Model

Recent developments concerning atomic structure shows that the Bohr picture of the atom with sharply defined electron orbits is not correct. According to the new theory of dual nature of matter, the electron in an atom does not behave as through it were a particle. The electrons move around the nucleus as *wave packets*, which are formed in a somewhat similar manner as standing waves are formed in sound.

According to this model electrons in various orbits behave as matter waves of wavelength $\lambda = h/mv$. The energy level and *orbits* of Bohr atom model are retained. The electron exists as a standing wave in each orbit. In fact, Bohr's quantum conditions which are assumed arbitrarily can be derived in a more natural way by considering the wave like property of the electron. For example, consider the electron in the hydrogen atom as a *standing wave* extending in a circle round the nucleus (Fig. 15.8). In order that this wave may just occupy the circumference of a circle, the circle must contain an integral number of wavelengths. If r is the radius of the circular orbit then

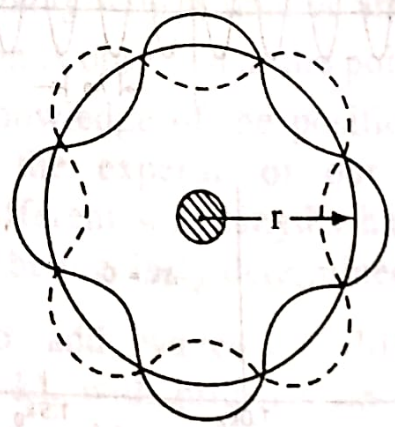


Fig. 15.8

$2\pi r = n\lambda$ where n is an integer (= 1,2,3,...)

$$\text{But } \lambda = \frac{h}{mv}; \quad \therefore 2\pi r = n\frac{h}{mv}$$

$$\text{or } mvr = n\frac{h}{2\pi}$$

But mvr is the angular momentum of the electron as a particle. Thus the wave-mechanical picture leads naturally to Bohr's postulate that the angular momentum is equal to an integral multiple of $h/2\pi$.

15.6 Heisenberg's Uncertainty Principle

As we have just seen, the evidence that matter is wavelike is very strong. At the same time, we cannot forget that the evidence that matter is particle-like is just as strong. The basic difference between these two view-points is that the position of a particle can be localized in both space and time but a wave cannot, being spread out in both of these dimensions. Let us explore the possibility of reconciling these two approaches by seeing whether we can put together an assembly of waves in such a way that we end up with something that reminds us of a particle. The conclusion that will be drawn must hold for all kinds of waves, whether they be water waves, sound waves, electromagnetic waves, or de Broglie waves. Let us start by trying to localize a wave in space and in time.

Localizing a wave in space. Let us consider a wave of the form $y = y_1 \sin k_1 x$ as shown in Fig.15.9. This is a wave that repeats itself endlessly from $x = -\infty$ to $x = +\infty$. If a question is asked "where is the wave located?", we cannot provide an answer – it is everywhere.

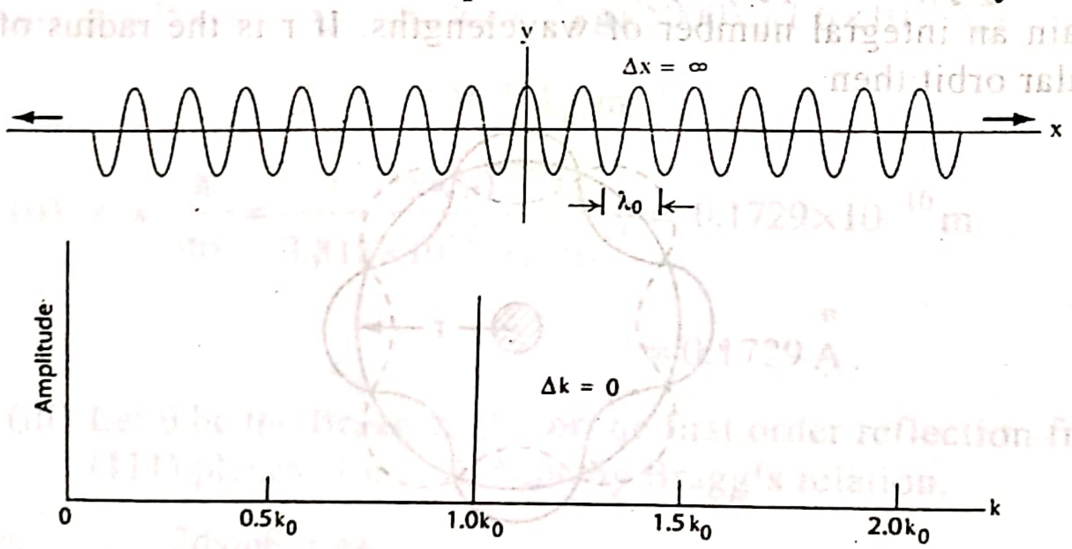


Fig. 15.9

Its wavelength, on the other hand is precisely determined and is equal to $2\pi/k_1$. If a wave is to be used to represent a particle, the wave must have one of the important attributes of a particle - it must be localized, or able to be confined to a relatively small (atom-sized or nucleus sized, for example) region of space. Because of its infinite extent, the pure sine wave is of no use in localizing particles.

Let us now see what happens when another wave of slightly different wavelength (*i.e*, different k) is added to our original wave, so that $y = y_1 \sin k_1 x + y_2 \sin k_2 x$. The characteristic pattern, known as *beats* in the case of sound waves, is produced as shown in Fig.15.10. The pattern still repeats endlessly from $x = -\infty$ to $x = +\infty$,

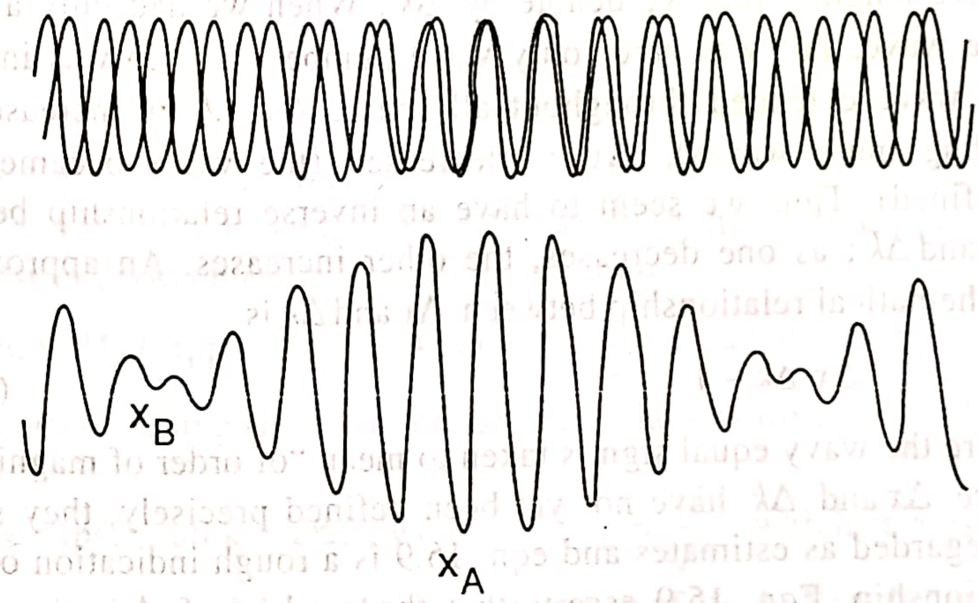


Fig. 15.10

but now a little bit more is known about the *position* of the wave – at certain values of x the medium is less likely to be waving than at others, or at least it is *waving* with a smaller amplitude. As can be seen in Fig.15.10, vibration is observed at the point $x = x_A$, but not $x = x_B$. The state of our knowledge of the position of the wave has improved, but it is at the expense of our knowledge of its wavelength. Since two different wavelengths have been added, the wavelength can no longer be precisely determined.

If we continue to add waves of different wavelengths (different wave numbers k), with properly chosen wave numbers, amplitudes and phases, we could eventually reach a situation similar to that illustrated in Fig. 15.11. This wave has virtually no amplitude

outside a rather narrow region of space Δx. Δx is not precisely determined, but is a rough measure of the region over which the wave has reasonably large amplitude. In order to achieve this situation, we must add together a large number of waves of different

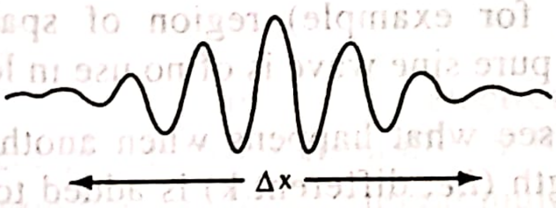


Fig. 15.11

wave numbers k. The wave thus represents a range of wave numbers (wave lengths) that we denote by Δk. When we had only a single sine wave, Δk was zero (only wave number) but Δx was infinite– the wave extended throughout all space. As Δk is increased (by adding more waves), Δx is decreased (the wave became more confined). Thus we seem to have an inverse relationship between Δx and Δk; as one decreases, the other increases. An approximate mathematical relationship between Δx and Δk is

$$\Delta x \, \Delta k \sim 1 \tag{15.9}$$

where the wavy equal sign is taken to mean "of order of magnitude". Since Δx and Δk have not yet been defined precisely, they should be regarded as estimates and eqn. 15.9 is a rough indication of their relationship. Eqn. 15.9 asserts that the product of Δx, the spatial extent of the wave, and Δk the range of wave numbers it contains, is of order of magnitude one.

For any type of wave, the position can be determined only at the expense of our knowledge of its wavelength. This statement, and its mathematical representation given by eqn. 15.9 are the first of our "uncertainty relationships" for classical waves. The position and the wavelength (wave number) are mutually "uncertain" to the degree given by eqn. 15.9.

Localizing a wave in time. A particle is localized in time as well as in space. If the space variable x in Fig. 15.9 is replaced by the time variable t (and the wavelength λ_0 by the time period T_0), that figure would then show how our wave would vary with time as it passes a particular fixed point, say, $x = 0$. As before, there is nothing at all

about this wave that suggests the localization in time that is associated with the word 'particle', because a particle would pass our observation point at a particular time, rather than spread over an infinite time interval.

Now a wave packet can be built up in time as well as in space. Fig. 15.11 can illustrate this, provided the space variables are replaced by the corresponding time variables, as above, and the wave number k_0 is also replaced by the angular frequency w_0. By analogy with eqn. 15.9, the duration Δt of our new wave packet is related to the spread Δw of angular frequencies needed to make up the wave packet by

$$\Delta w.\Delta t \sim 1 \tag{15.10}$$

This is the second of the uncertainty relationships, for classical waves, and is similar to eqn. 15.9 ($\Delta x.\Delta k \sim 1$) in that it gives a rough relationship between estimates of uncertainties.

Eqn. 15.9 applies to all kinds of waves. Let us apply to de Broglie wave. Using the basic de Broglie relationship $p = \dfrac{h}{\lambda}$ along with the expression $k = 2\pi/\lambda$, we find that $p = \dfrac{hk}{2\pi}$, which relates the momentum of a particle to the wave number of its de Broglie wave. The combination $h/2\pi$ occurs frequently in wave mechanics and is given the special symbol $\hbar$ ("h-bar")

$$\hbar = \frac{h}{2\pi} = 1.05\times10^{-34}\,\text{J.s} = 6.58\times10^{-16}\,\text{eV.s}$$

In terms of $\hbar$,

$$p = \hbar k$$

$$\therefore \quad \Delta k = \frac{\Delta p}{\hbar}$$

Hence from the relation (eqn. 15.9) we have

$$\Delta x.\Delta p_x \sim \hbar \tag{15.10a}$$

The x subscript has been added to the momentum to remind us that eqn. 15.10a applies to motion in a given direction and relates the uncertainties in position and momentum in that direction only. Similar and independent relationships can be applied in the other directions as necessary. Thus

$$\Delta x.\Delta p_x \sim \hbar$$

$$\Delta y.\Delta p_y \sim \hbar \qquad (15.11)$$

$$\Delta z.\Delta p_z \sim \hbar$$

Equations 15.11 are the *Heisenberg uncertainty relationships* first derived by Werner Heisenberg in 1927. They are the mathematical formulation of *Heisenberg's uncertainty pariciple.*

It is not possible to determine simultaneously both the position and momentum of a particle with unlimited precision.

The energy-frequency relationship for photons is E = hν, which can be written as $E = h.\frac{w}{2\pi} = \hbar.w$. As in the case of the momentum-wavelength relationship, this relationship can be carried over to particles for the purpose of calculating the uncertainty in energy, which involves small differentials only. Thus we can write $\Delta E = \hbar.\Delta w$.

By substituting for Δw in eqn. 15.12 we have,

$$\Delta E\,.\,\Delta t \sim \hbar \qquad (15.12)$$

or,

It is not possible to make a simultaneous determination of the energy and the time coordinate of a particle with unlimited precision.

According to classical ideas, it is possible for a particle to occupy a fixed position and have a definite momentum and we can predict *exactly* its position and momentum at any time later. But according to uncertainty principle, it is not possible to determine accurately the simultaneous values of position and momentum of a particle at any time. Eqn. 15.11($\Delta x.\Delta p_x \sim \hbar$) gives an estimate of

the minimum uncertainty that can result from any experiment; measurement of the position and momentum of a particle will give a spread of values of widths Δx and Δp_x. We may, for other reasons, do much worse than eqn.15.11, but we can do no better. Occasionally eqns. 15.11 and 15.12 are written with $\frac{\hbar}{2}$ or h rather than $\hbar$, on the right hand side, or else with > rather than ~ showing the equality. This difference is not very important, since eqns. 15.11 and 15.12 give us only estimates. The actual uncertainties Δx and Δp_x depend on the distribution of wave number (or wavelengths) used to restrict the wave to the region Δx. The most compact distribution $\Delta x \,.\, \Delta p_x > \frac{\hbar}{2}$ while for all other distributions $\Delta x \,.\, \Delta p_x > \frac{\hbar}{2}$. We are therefore safe in using $\hbar$ as an estimate.

Uncertainty principle has a profound impact on our view of nature. Eqns. 15.11 and 15.12 imply that *nature imposes a limit on the accuracy with which we can do experiments*; no matter how well designed our measuring apparatus might be; we still can do no better than eqns 15.11 and 15.12 To determine the momentum accurately, measurements must be taken over a long distance Δx; if the particle is to be confined to a small region of space Δx; the ability to measure its momentum is lost. To measure an energy with small uncertainty takes a long time Δt; if a particle lives for a short time, its energy uncertainty will be large.

The uncertainty principle does *not* limit our precision of measurement when we are dealing with large objects such as golf balls. Here ordinary instrumental errors overwhelm the fundamental limits set by this principle. When we deal with electrons and other elementary particles, however, the situation is quite different.

The following examples give applications of the uncertainty relationships.

(i) Suppose we try to measure the position and linear momentum of an electron using an imaginary microscope with high resolving power (Fig. 15.12). Since the lower limit of resolution depends upon the wavelength of the light employed to illuminate the

particle, it follows that radiation of shortest wavelength, such as γ-rays, must be used, if we wish to determine the position as accurately as possible. According to the corpuscular view, the minimum amount of light that could be used for irradiation is a single quantum hν. The electron can be observed if at least one photon is scattered by it into the microscope lens. The resolving power of the microscope is given by the relation

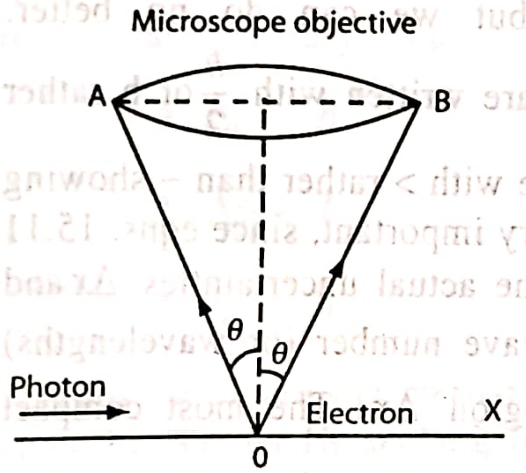


Fig. 15.12

$$\Delta x = \frac{\lambda}{2\sin\theta}$$

where Δx is the distance between two points which can be just resolved by the microscope. This is the range in which the electron would be visible when irradiated by the photon. Hence Δx is the uncertainty involved in the position measurement of the electron.

However, the incoming photon will interact with the electron through the Compton effect. The electron suffers a Compton recoil of the order of magnitude $\frac{h\nu}{c}$. To able to see the electron, the scattered photon should enter the microscope within the angle 2θ. The component of the momentum of the electron along OA is $-\frac{h\nu}{c}\sin\theta$ and that along OB is $\frac{h\nu}{c}\sin\theta$. Hence the uncertainty in the momentum measurement in the x-direction is

$$\Delta p_x = \frac{h\nu}{c}\sin\theta - \left(-\frac{h\nu}{c}\sin\theta\right) = \frac{2h\nu}{c}\sin\theta = \frac{2h}{\lambda}\sin\theta$$

$$\therefore \quad \Delta x \times \Delta p_x = \frac{\lambda}{2\sin\theta} \times \frac{2h\sin\theta}{\lambda} = h$$ (Heisenberg's uncertainty principle)

A more sophisticated approach will show that $\Delta x . \Delta p_x \geq \frac{h}{2\pi}$.

(ii) Diffraction of a beam of electrons by a slit:

Let us consider a beam of electrons of speed v_0, moving upward as in Fig. 15.13. We would like to measure simultaneously and with unlimited precision the horizontal position x and the velocity component v_x for an electron in this beam. As we shall see, this cannot be accomplished.

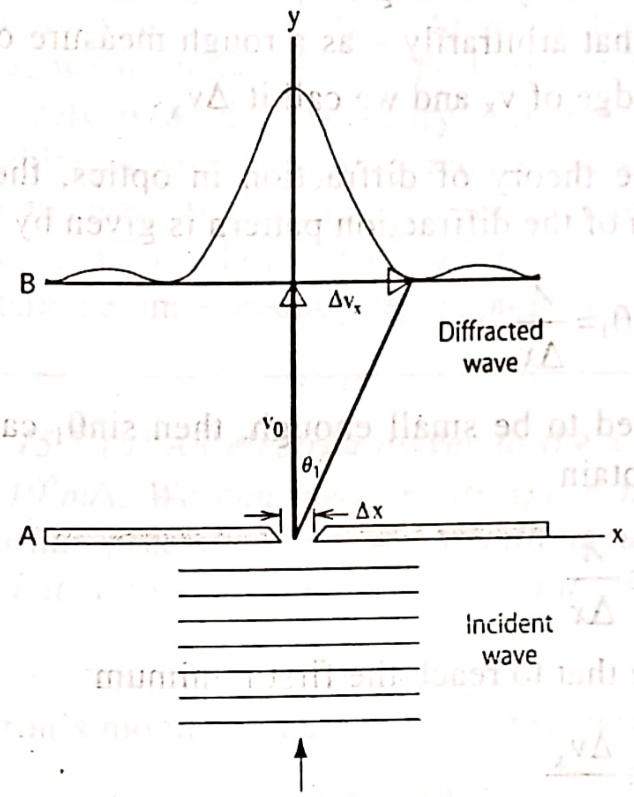


Fig. 15.13

To measure x let us block the beam with a screen A in which we put a slit of width Δx. If an electron passes through this slit, we can claim to know its horizontal position to this precision. By narrowing the slit, the precision of this measurement can be improved as much as we wish.

However, something else happens that perhaps we have not counted upon. The electron beam, being wavelike, flares out by diffraction as it passes through the slit. If a suitably sensitive screen B is placed as in Fig. 15.13 a typical single-slit diffraction pattern show up. Electrons that form left half of the pattern, must have been moving to the left - some faster, some slower - as they emerged from the slit. Those that form the right half of the pattern, must have

been moving to the right. Even though the symmetry of the arrangements requires that the average value of the emerging electrons is zero, individual electrons can have nonzero values.

Let there be a particular value of v_x that will cause the electron to land at the position of the first minimum of the diffraction pattern, identified by the angle θ_1 in Fig. 15.13. We take this value of v_x –somewhat arbitrarily – as a rough measure of the uncertainty of our knowledge of v_x and we call it Δv_x.

From the theory of diffraction in optics, the location of the first minimum of the diffraction pattern is given by

$$\sin\theta_1 = \frac{\lambda}{\Delta x}$$

If θ_1 is assumed to be small enough, then $\sin\theta_1$ can be replaced by θ_1. We then obtain

$$\theta_1 \cong \frac{\lambda}{\Delta x}$$

It must be true that to reach the first minimum

$$\theta_1 \cong \frac{\Delta v_x}{v_0}$$

Combining these two relations we obtain

$$\Delta v_x . \Delta x \cong \lambda v_0$$

Now λ, the de Broglie wavelength of the electron is equal to h/mv_0. Putting this into the above relation we get

$$\Delta v_x . \Delta x \cong \frac{hv_0}{mv_0} \equiv \frac{h}{m}$$

$$\text{or,} \quad m.\Delta v_x . \Delta x \equiv h$$

$$\text{or,} \quad \Delta p_x \Delta x \equiv h$$

This is certainly consistent with eqn. 15.11; minor difference (the factor of 2π) result from the arbitrary way of defining Δp_x and Δx.

We now have a clear picture of how uncertainty principle operates in this case. If we want to pin down the horizontal position of the electron, we must narrow the slit. This however broadens the diffraction pattern so that Δp_x increases. If on the other hand, we want to pin down the horizontal component of the momentum of the electron, the angular width of the diffraction pattern must somehow be reduced. The only way to do this is to widen the slit but that, in turn, means that we no longer know the horizontal position of the electron as precisely as we did. As we try to increase our knowledge about one variable, we simultaneously reduce our knowledge about the other. The uncertainty principle is not a statement about the electron or any other particle for that matter, it is a statement about our ability to determine simultaneously certain properties of those particles.

Example 15.7 (a) An electron moves in the x direction with a speed of 3.6× 10⁶m/s. We can measure its speed to a precision of 1%. (i) With what precision can we simultaneously measure its position? (b) What can we say about its motion in the y direction?

Soln.

The electron's momentum along the x direction is

$$p_x = mv_x = (9.1\times10^{-31}\text{kg})(3.6\times10^{6}\text{m/s})$$

$$= 3.3\times10^{-24}\ \text{kg.m/s.}$$

The uncertainty Δp_x is 1% of this value and is equal to

$$= (3.3\times10^{-24})(0.01)\ \text{kg.m/s}$$

$$= 3.3\times10^{-26}\ \text{kg.m/s}$$

The uncertainty in its position is then given by the relation

$$\Delta p_x . \Delta x \equiv \hbar$$

$$\text{or } \Delta x \equiv \frac{\hbar}{\Delta p_x} = \frac{6.62\times10^{-34}\ \text{J.s}}{(2\times3.141)(3.3\times10^{-26})\text{kg.m/s}}$$

$$= 3.2\times10^{-9}\ \text{m} = 3.2\ \text{nm.}$$

This is roughly equal to 10 atomic diameters.

(b) If the electron moves in the x direction, then its speed in the y direction is known precisely;

or $\Delta p_y = 0$. Then the uncertainty relationship

$\Delta p_y . \Delta y \equiv \hbar$ demands that

$$\Delta y \equiv \frac{\hbar}{0} \equiv \infty \text{ (infinity)}.$$

Thus we cannot know anything at all about the electron's y-coordinate.

***Example 15.8** Consider a pitched baseball (m = 0.145 kg) to be moving with a speed of 42.5 m/s. Again it is assumed that the speed can be measured to a precision of 1%. With what precision can we simultaneously measure its position?*

Soln.

The momentum of the baseball

$$p_x = mv_x = (0.145\text{kg})(42.5\text{m/s})$$

$$= 6.16 \text{ kg.m/s.}$$

The uncertainty in the momentum is

$$(6.16 \text{ kg.m/s})(0.01) = 6.16 \times 10^{-2} \text{ kg.m/s.}$$

The corresponding uncertainty in its position is

$$\Delta x \equiv \frac{\hbar}{\Delta p_x} \equiv \frac{1.05 \times 10^{-34} \text{ J.s}}{6.16 \times 10^{-2} \text{kg.m/s}}$$

$$= 1.7 \times 10^{-33} \text{ m.}$$

This uncertainty is 19 orders of magnitude smaller than the size of an atomic nucleus.

***Example 15.9** A microscope, using photons, is employed to locate an electron in an atom to within a distance of 0.2 Å. What is the uncertainty in the momentum of the electron located in this way?*

Soln.

$$\Delta x = 0.2\,\overset{o}{A} = 0.2\times10^{-10}\,m$$

From the uncertainty relation $\Delta x.\Delta p = \frac{h}{2\pi}$

We get

$$\Delta p \equiv \frac{\hbar}{\Delta x} = \frac{6.62\times10^{-34}\,J.s}{(2\times3.141)(0.2\times10^{-10}\,m)}$$

$$= 5.274\times10^{-24}\ kg.m/s.$$

Example 15.10 *The lifetime of an excited state of an atom is about 10^{-8} sec. Calculate the minimum uncertainty in the determination of the energy of the excited state.*

Soln.

The uncertainty relation connecting time and energy is given by

$$\Delta E.\Delta t \equiv \hbar$$

Here $\Delta t = 10^{-8}$ s.

$$\therefore \quad \Delta E \equiv \frac{h}{2\pi.\Delta t} = \frac{6.62\times10^{-34}\,J.s}{(2\times3.141)(10^{-8}\,s)} = 1.0\times10^{-26}\ J$$

$$= 6.5\times10^{-8}eV.$$

This is known as the *energy width* of an excited state.

Example 15.11 *In nuclear beta decay, electrons are observed to be ejected from the atomic nucleus. Suppose we assume that electrons are somehow trapped within the nucleus, and that occasionally one escapes and is observed in the laboratory. Take the diameter of a typical nucleus to be 1.0×10^{-14} m and use the uncertainty principle to decide if electrons can exist within the nucleus.*

Soln.

The electron is supposed to be trapped in a region of width $\Delta x \sim 10^{-14}$m. The corresponding uncertainty in its momentum will be

$$\Delta p_x = \frac{\hbar}{\Delta x} = \frac{1}{c}\frac{\hbar c}{\Delta x} = \frac{1}{c}\frac{197\text{MeV.fm}}{10\text{fm}} = 19.\frac{\text{MeV}}{c}$$

We have used $\hbar c = 197\text{MeV.fm}$ in this calculation. The momentum is clearly in the relativistic region for electrons. We must therefore use the relativistic formula to find the kinetic energy.

$$K^2 = p^2c^2 + m_0^2c^4$$

As the rest mass energy of electron is of the order of 0.511 MeV, which is much smaller than the value of the first term, it can be neglected. Then

$$K^2 = p^2c^2$$

Now the momentum component p_x and hence the magnitude of the total momentum p of the electron in the nucleus can be related by the equation

$$\Delta p_x = \sqrt{(p_x)^2_{av}}$$

This gives us

$$K = 19 \text{ MeV}.$$

Thus, if the electrons are to exist inside the nucleus, they must have a spread of kinetic energies with a typical value of around 19 MeV. However, we know that the electrons emitted by radioactive nuclei during beta decay have energies of only 3 to 4 MeV. So, in general electrons cannot exist within the nucleus. Hence, another explanation must be found for electrons observed in nuclear beta decay.

15.7 Electron microscope

A particle of mass m moving with a large velocity v behaves like a wave of wavelength $\lambda = h/mv$. This wave nature of the electron has found wide practical applications, the most important of which is the invention of the electron microscope. Microscopes are used not only to produce magnified image of very small objects but also to reveal detailed structures of such objects. Higher the

resolving power of a microscope, more minutely can the detailed structure of an object be studied with the microscope.

In an optical microscope, the object is illuminated with visible light. The limit of resolution of a microscope depends on the wavelength of light used and is given by

$$\Delta x = \frac{0.61\lambda}{\text{numeriacl aperture}}$$

where Δx is the minimum separation between two point objects which can be just resolved. The numerical aperture depends on the angle subtended by the objective lens at the object and the refractive angle of the object space. Smaller the Δx, the better is the limit of resolution. This means that, for better resolution, the object should be illuminated with light of the shortest possible wavelength. Thus in an optical microscope the limit of resolution is of the order of the wavelength of visible radiation which is a few thousand angstroms. As the de Broglie wavelength is much shorter than that of visible radiation, a much better resolution can be obtained if the object is illuminated with electron waves. Thus for electrons accelerated by a potential difference of 10^5 volts, the energy of the electron beam is 10^5eV and the de Broglie wave length associated with the electron is

$$\lambda = \frac{ch}{\sqrt{eV(eV + 2m_0c^2)}}$$

$$= \frac{(3\times10^8)(6.62\times10^{-34})}{[10^5(10^5 + 2\times0.51\times10^6)]^{1/2}\times(1.6\times10^{-19})}$$

$$= 0.0371 \times 10^{-10}\text{m}$$

$$= 0.0371\ \overset{o}{A}.$$

With higher energy electrons, the wavelength is even shorter. So it is possible to construct electron microscopes with a limit of resolution of the order of a fraction of an angstrom quite easily.

The schematic diagram of the different parts of an electron microscope is shown in Fig. 15.14. In place of the lens system used for the different parts in an optical microscope, electric and magnetic fields

are used in an electron microscope. It is possible to focus the electron beam with the help of the magnetic field produced by a current carrying coil which thus acts as a lens in an electron microscope.

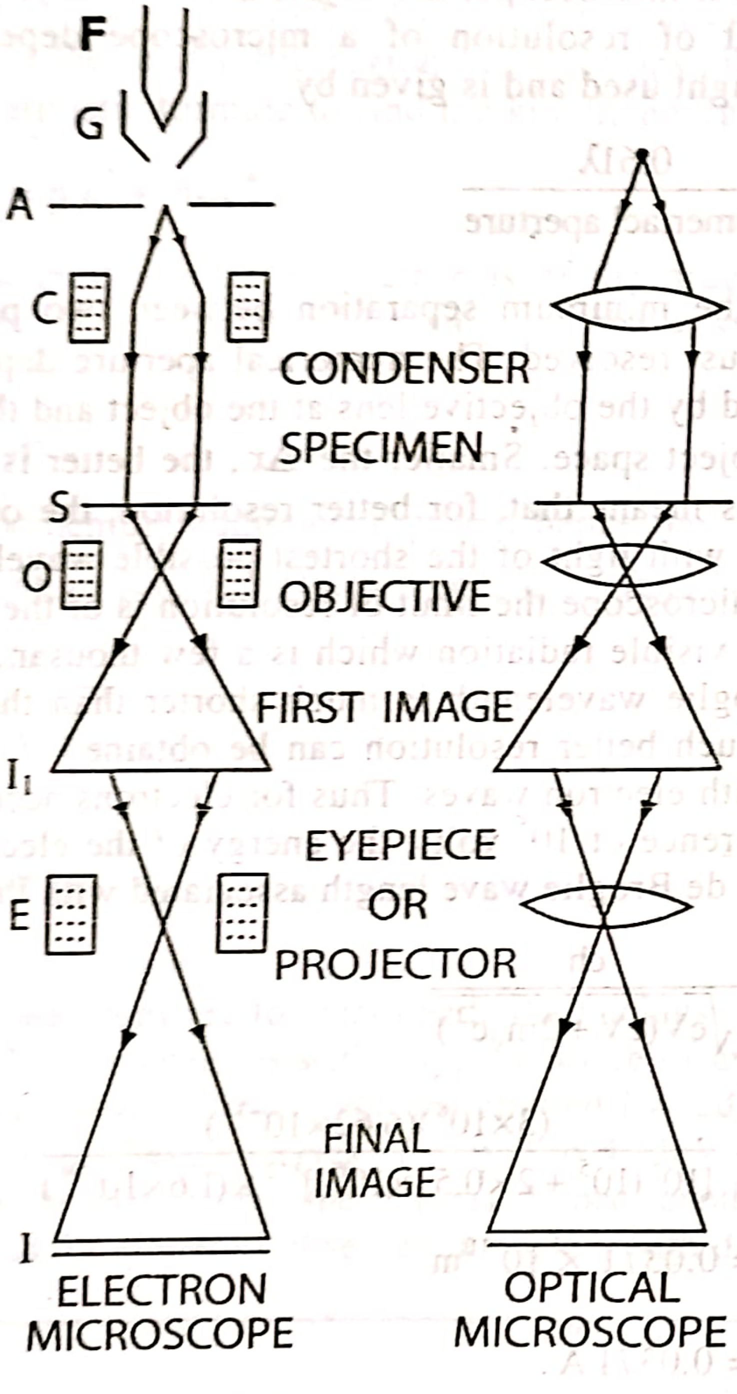


Fig. 15.14

In Fig. 15.14 electrons, liberated by thermionic emission from a heated filament F, are first accelerated through a large potential difference (~50 to 100 kilo-volts) towards the anode A. Before reaching the anode, the electron pass through the grid G. By adjusting the grid potential, the intensity of the electron beam can be

controlled. The accelerated electron beam passes through a small hole in the anode and is condensed with the help of the magnetic field. The magnetic field is generally provided by two current carrying coils (c) which are enclosed in two iron cases having one hole each. When the holes face each other, the magnetic field in the space between the two coils produces convergence. The focal length of such a magnetic lens depends upon the quality of the coil and the current it carries. The intensity of the magnetic field can be adjusted by adjusting the currents in the coils and thus the focal length of the system (magnetic lens) can be adjusted.

The electron beam so condensed by the condenser C falls on the specimen S to illuminate it. The specimen must be very thin so that the illuminating electron beam may pass through it without losing any appreciable energy. The emergent beam then passes through another magnetic field produced by the coil O which acts as the objective lens. The objective lens produces a magnified intermediate image I_1. The electrons emerging from I_1 then pass through another magnetic field produced by the coil E which acts as the eye-piece. The final image I formed by the eye-piece is either projected on a fluorescent screen or can be photographed by a suitable arrangement.

Fig. 15.15 shows side by side the corresponding parts of an optical microscope and an electron microscope for comparison.

It is to be noted that the pressure within the electron microscope must be kept very low ($<10^{-5}$mm of Hg) so that the electron beam may not suffer scattering or absorption within the microscope.

Uses

(1) Electron microscope is used in industry to study the structures of textile fibres, surface of metals, composition of paper and paints, etc.

(2) In the field of medicine and biology, electron microscope is used to study virus, a disease causing agent, which is beyond the scope of ordinary microscope. A knowledge of the structure of viruses enables scientists to devise methods of their destruction and thereby control diseases. Bacteria are also shown in greater detail in an electron microscope.

(3) In physics, electron microscope is used to investigate the structure of crystals and atomic structure.

(4) Because of electron microscope, small particles forming colloids in chemistry can now be studied.

15.8 Wave Mechanics

Certain basic postulates are of fundamental importance in the development of Wave Mechanics. The postulates are three in number. These are

(1) *Each dynamical variable relating to the motion of a particle can be represented by a linear operator.*

In classical physics, certain definite functions of suitable variables are associated with each observable quantity. For example (x,y,z) are associated with position, mv with momentum, $\frac{1}{2}mv^2$ with kinetic energy and so on. Similarly, in wave mechanics and quantum mechanics, certain operators are associated with observable quantities. The x-component of the linear momentum of a particle has a classical expression $p_x = m\left(\frac{dx}{dt}\right)$. For angular momentum the operator can be written as $r \times p = -i(\frac{h}{2\pi})(r \times \nabla)$. Similarly, the classical representation for the observed *total energy* is $\frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x,y,z)$ and the quantum mechanical operator is

$$-\frac{(h/2\pi)^2}{2m}\left(\frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2} + \frac{\delta^2}{\delta z^2}\right) + V(x,y,z)$$

An operator tells us what operation to carry out on the quantity that follows it. Thus, the operator $i(\frac{h}{2\pi})\frac{\delta}{\delta t}$ instructs us to take the partial derivative of what comes after it with respect to t and multiply the result by $i(\frac{h}{2\pi})$.

The quantum operators for several physical quantities are summarized below:

Quantity	Classical definition	Quantum operator
Position	r	r
Momentum	p	$-i\frac{h}{2\pi}\nabla$
Angular momentum	$r \times p$	$-i\frac{h}{2\pi} r \times \nabla$
Kinetic energy	$p^2/2m$	$-(h^2/8\pi^2 m)\nabla^2$
Total energy	$p^2/2m + E_p(\underline{r})$	$-(h^2/8\pi^2 m)\nabla^2 + E_p(\underline{r})$

(2) *A linear eigenvalue equation can always be linked with operator.*

The total energy operator is $i(\frac{h}{2\pi})\frac{\delta}{\delta t}$. Let us consider the eigenvalue equation $i(\frac{h}{2\pi})\frac{\delta\psi}{\delta t} = E\psi$. Here ψ is said to be the eigen function of the operator $i(\frac{h}{2\pi})\frac{\delta}{\delta t}$ and E is called the corresponding energy eigenvalue.

(3) When a measurement of a dynamical quantity a is made on a particle for which the wave function is Ψ, one generally gets different values of a during different trials. This is in conformity with the uncertainty principle. The most *probable value* of a is given by

$$< a > = \int_0^\infty \psi^* \hat{A} \psi dV$$

where $\hat{A}$ is the operator associated with the quantity a and Ψ^* is the complex conjugate of Ψ. The quantity $< a >$ is called the expectation

value of $\hat{A}$ and is defined as the value of a obtained in the majority of the trials. For example, the average momentum is given by

$$<p_x> = \int \psi^* (-i\frac{h}{2\pi}\frac{\delta}{\delta x})\psi dx$$

15.9 Schrödinger's fundamental wave equation

To analyze the wave behavior of the electron, we require a mathematical procedure in which we can specify the interaction of the electron with its environment and then solve for its motion. This is of course just what was done in classical physics using Newton's laws, in which the interaction were described in terms of forces.

The wave-mechanical procedure for studying the behavior of electrons (and other particles) is based on an equation proposed by the Austrian physicist Erwin Schrödinger. In 1926, just 2 years after de Broglie's hypothesis concerning matter waves, Schrödinger directly started with de Broglie's idea of matter waves and developed it into a rigorous mathematical theory which is now known as wave mechanics. The essential feature of this theory is the incorporation of the expression for the de Broglie wavelength into the general classical wave equation. This led to the derivation of a wave equation for a moving particle, which is known as *Schrödinger's fundamental wave equation.*

According to the de Broglie theory, a particle of mass m and moving with a velocity v has a wave system of some kind associated with it, the wavelength λ being given by λ=h/mv. Although we have no precise knowledge about the wave system, we shall assume that it can be indicated by ψ, the periodic changes in which gives rise to the wave system.

Let the position of the particle be located at a point x,y,z in the Cartesian system of co-ordinates. Let ψ, in the immediate vicinity of the particle undergo periodic changes, its value at any instant t being given by

$$\psi = \psi_0 \sin 2\pi \nu t \qquad \text{(i)}$$

where ψ_0 is the amplitude at the point considered, independent of t but a function of x,y,z and ν the frequency.

The differential equation of this wave equation can be written in the form

$$\frac{\delta^2\psi}{\delta t^2} = v^2\left[\frac{\delta^2\psi}{\delta x^2}+\frac{\delta^2\psi}{\delta y^2}+\frac{\delta^2\psi}{\delta z^2}\right] \quad \text{(ii)}$$

$$= v^2\nabla^2\psi$$

where $\nabla^2\psi = \frac{\delta^2\psi}{\delta x^2}+\frac{\delta^2\psi}{\delta y^2}+\frac{\delta^2\psi}{\delta z^2}$, ∇^2 being the Laplacian operator.

From eqn (i),

$$\frac{\delta^2\psi}{\delta t^2} = -4\pi^2\nu^2\psi = -4\pi^2\frac{v^2}{\lambda^2}\psi$$

since $\nu = v/\lambda$.

Substituting this in eqn. (ii),

$$-\frac{4\pi^2 v^2\psi}{\lambda^2} = v^2\nabla^2\psi$$

$$\text{or} \quad \nabla^2\psi + \frac{4\pi^2}{\lambda^2}\psi = 0 \quad \text{(iii)}$$

So far the treatment is general. Let us now introduce the wave mechanical concept and replace λ by h/mv from de Broglie's theory. With this change eqn. (iii) becomes

$$\nabla^2\psi + \frac{4\pi^2 m^2 v^2}{h^2}\psi = 0 \quad \text{(iv)}$$

Now the total energy of a particle is

E = Kinetic energy + Potential energy

or, Kinetic energy, $\frac{1}{2}mv^2 = E - V$

or, $m^2v^2 = 2m(E - V)$

where V is the potential energy of the particle.

Substituting this in eqn. (iv), we obtain

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0 \qquad (15.13)$$

This is the time-independent form of the Schrödinger wave equation for a simple particle.

Eqn. (15.13) may be written as

$$\left(\frac{-(h/2\pi)^2}{2m}\nabla^2 + V\right)\psi = E\psi$$

An operator *H*, called the Hamiltonian operator, can be defined as

$$H = -\frac{(h/2\pi)^2}{2m}\nabla^2 + V$$

So, in operator form, the time independent Schrödinger equation can be written as

$$\hat{H}\psi = E\psi \qquad (15.14)$$

The function cannot be cancelled out of this equation, because H is not a simple scalar multiplier, while E is the value of an energy. Eqn(15.14) must be interpreted as

(operator H) acting on function ψ = (total energy) multiplying function ψ.

Schrödinger equation can be applied for determining the total energy of an electron when it is moving in an electric field so that its potential energy V is specified in terms of space coordinates and not of time. Like other differential equations, its solutions are also governed by boundary conditions. Generally solutions are only obtained for certain energy values called *characteristic* or *eigenvalues*. The corresponding wave function ψ are called *characteristic* or *eigen-functions*.

Physical significance of ψ

The probability that a particle will be found at a given place in space at a given instant of time is characterized by the function ψ (x,

y, z, t). It is called the wave function and can be either real or complex. If we write $\psi = a + jb$ and its conjugate $\psi^* = a - jb$, then their product $\psi\psi^* = a^2 + b^2$ and is real. The quantity $P = |\psi|^2 = \psi\psi^*$ is the probability density. For a one-dimension system, $|\psi|^2 dx$ represents the probability of finding the electron within the range x to (x + dx). But in three dimensions, $|\psi|^2 dV$ represents the probability of finding an electron within the volume range V to (V+dV). Further, since the particle is certainly to be found somewhere in space

$$\int |\psi|^2 dV = 1$$

or $$\iiint |\psi|^2 dx\,dy\,dz = 1 \tag{15.15}$$

the triple integral extending over all possible values of x, y, z. A wave function $|\psi|$ satisfying this relation is called a *normalized wave function*.

Orthogonal and normalized wave functions

If the product of a function $\psi_1(x)$ and the complex conjugate $\psi_2^*(x)$ of a function $\psi_2(x)$ vanishes when integrated with respect to x over the interval $a \le x \le b$, *i.e.*, if

$$\int_a^b \psi_2^*(x)\psi_1(x)dx = 0$$

then $\psi_1(x)$ and $\psi_2(x)$ are said to be orthogonal in the interval (a,b).

We already know that the probability of finding a particle within a volume element is $|\psi|^2 dV$. The total probability of finding the particle in the entire space is, of course, unity, *i.e*,

$$\int |\psi|^2 dV = 1$$

where the integration extends over the entire space. The above equation can also be written as

$$\int \psi\psi^* dV = 1$$

Any wave function satisfying the above equation is said to be normalized to unity or simply normalized.

Very often ψ is not a normalized wave function. However, it is possible to multiply ψ by a constant A, to give a new wave function $A\psi$, which is a solution of the wave equation. The problem is to choose the proper value of A so that the new wave function is a normalised wave function. For it to be a normalised wave function, it must meet the requirement

$$\int (A\psi)^*(A\psi)dxdydz = 1$$

or $$|A|^2 \int \psi\psi^* dxdydz = 1$$

or $$|A|^2 = \frac{1}{\int \psi\psi^* dxdydz}$$

$|A|$ is known as the *normalizing* constant.

In order to arrive at results consistent with physical observation, several additional requirements are imposed on the wave function $\psi(x)$. These are

1. It must be single-valued and continuous everywhere.
2. If $\psi_1(x)$... ... ...$\psi_n(x)$ are solutions of Schrödinger equation, then the linear combination $\psi(x) = a_1\psi_1(x)+a_2\psi_2(x)+\ldots\ldots\ldots+a_n\psi_n(x)$ must also be a solution.
3. The wave function $\psi(x)$ must approach zero as $x \to \pm\infty$

Eigenfunctions and Eigenvalues

Schrödinger time-independent wave equation is an example of a type of differential equation called an *eigenvalue equation*. In general, an eigenvalue equation can be written as

$$F_{op}\psi = f\psi$$

The differential operator F_{op} operates on a function ψ, and this yields a constant f times the function. The function f is called the *eigenfunction* of the operator F_{op} and the corresponding value for f is called the *eigenvalue*.

Applications of Schrödinger equation

(a) Particle in a box

Consider a particle moving freely in a one-dimensional box of length L. The walls are assumed to be perfectly reflecting and the particle is bouncing back and forth between the walls of the box. The box has insurmountable potential barriers at x = 0 and x = L *i.e*, the box is supposed to have walls of infinite height at x = 0 and x = L. The potential energy V of the particle is infinite on both sides of the box and is shown in Fig. 15.15. This is known as an *infinite potential*

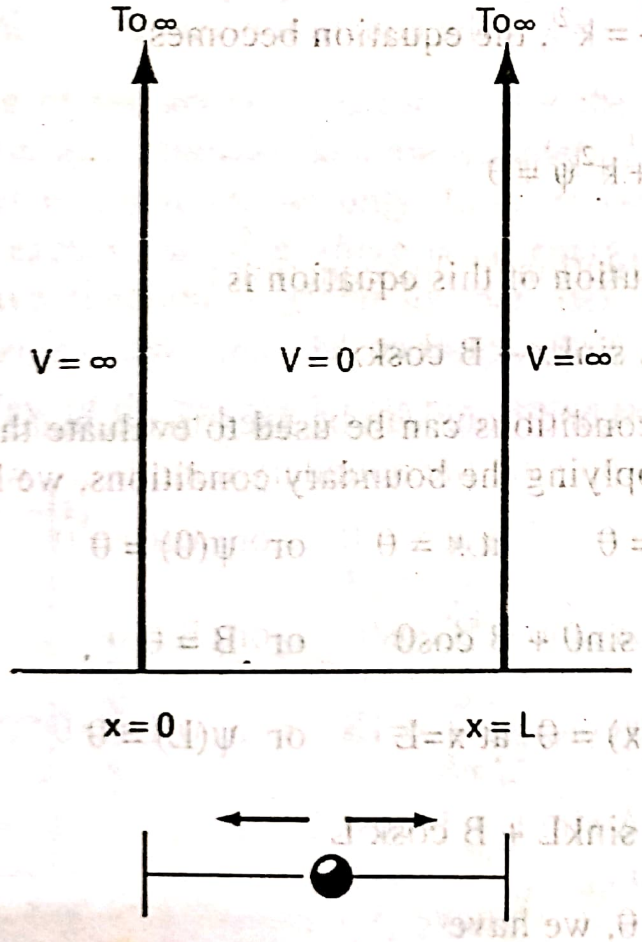


Fig. 15.15

energy well. Of course, we are free to choose any constant value for V in the region $0 \leq x \leq L$; we choose it to be zero for convenience. The potential energy of the particle may be expressed as

$$V(x) = 0 \qquad 0 \le x \le L$$
$$= \infty \qquad x < 0,\ x > L$$

If the walls of the box are perfectly rigid, the particle must always be in the box and the probability of finding it elsewhere must be zero. To make the probability zero everywhere outside the box, we must make $\psi = 0$ outside the box. Thus we have

$$\psi(x) = 0 \qquad x < 0,\ x > L$$

The Schrödinger equation for $0 \le x \le L$, when $V(x) = 0$ is.

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2} E\psi = 0$$

where m is the mass of the particle.

Putting $\frac{8\pi^2 mE}{h^2} = k^2$, the equation becomes

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0$$

The general solution of this equation is

$$\psi = A\ \text{sin}kx + B\ \text{cos}kx \qquad \text{(i)}$$

The boundary conditions can be used to evaluate the constant A and B in eqn (i). Applying the boundary conditions, we have

$$\psi(x) = 0 \qquad \text{at } x = 0 \qquad \text{or} \quad \psi(0) = 0$$

$$\therefore\ 0 = A\ \text{sin}0 + B\ \text{cos}0 \qquad \text{or} \quad B = 0$$

$$\text{Also} \quad \psi(x) = 0 \quad \text{at } x = L \qquad \text{or} \quad \psi(L) = 0$$

$$\therefore\ 0 = A\ \text{sin}kL + B\ \text{cos}kL$$

since B = 0, we have

$$0 = A\ \text{sin}kL$$

A is not zero; so sinkL must be equal to zero. It will be possible only when

$$kL = n\pi \quad \text{or} \quad k = \frac{n\pi}{L}$$

where n is called quantum numbers and equals 1,2,3,....∝

$$\text{Thus } \psi_n(x) = A\sin\frac{n\pi x}{L} \qquad \text{(ii)}$$

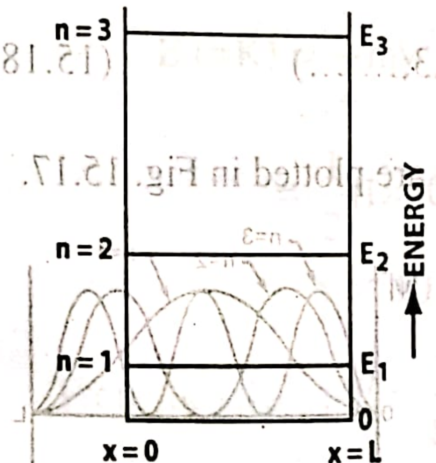
Substituting this value in eqn. (15.13) and also putting V = 0, we get

$$\frac{n^2\pi^2}{L^2} = \frac{8\pi^2 m}{h^2}.E$$

Here, E – the total energy of the electron – becomes equal to its kinetic energy.

$$\therefore \; E_n = \frac{n^2h^2}{8mL^2} \quad \text{where } n=1,2,3,......∝ \qquad (15.17)$$

The meaning of the above equation is that the particle in the box cannot possess any arbitrary amount of energy. Rather it can have energy values in a discrete set only. In other words, its energy is quantized. For each value of n, there is an energy level and the corresponding wave function is given by eqn. (ii). Each value of E_n is called an *eigenvalue* and the corresponding ψ_n is called *eigenfunction*. A few of the energy levels are shown in Fig. 15.16.



$$E_1 = \frac{h^2}{8mL}$$

$$E_2 = \frac{4h^2}{8mL^2}$$

$$E_3 = \frac{9h^2}{8mL^2} \quad \text{and so on}$$

It may be noted that the particle cannot have zero energy.

The particle in a box: Wave functions

Our solution of $\psi(x)$ is not yet complete, since we have not yet determined the constant A. To do this, we note that the particle is somewhere inside the box. We, therefore, go back to the normalization condition given by $\int_{-\infty}^{+\infty}\psi^2 dx = 1$. Since $\psi = 0$ except for $0 \le x \le L$, the integral vanishes except inside that region, sot that

$$\int_0^L \psi^* \psi dx = 1$$

i.e., $$\int_0^L A^2 \sin^2\left(\frac{n\pi x}{L}\right)dx = 1$$

or $$A^2 \int_0^L \left(\frac{1-\cos\frac{2n\pi x}{L}}{2}\right)dx = 1$$

or $$A^2 \frac{L}{2} = 1$$

or $$A = \sqrt{\frac{2}{L}}$$

The normalized wave function of the particle is then

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\frac{n\pi x}{L} \qquad (n = 1,2,3,\ldots\ldots) \qquad (15.18)$$

The normalized wave functions ψ_1, ψ_2 and ψ_3 are plotted in Fig. 15.17.

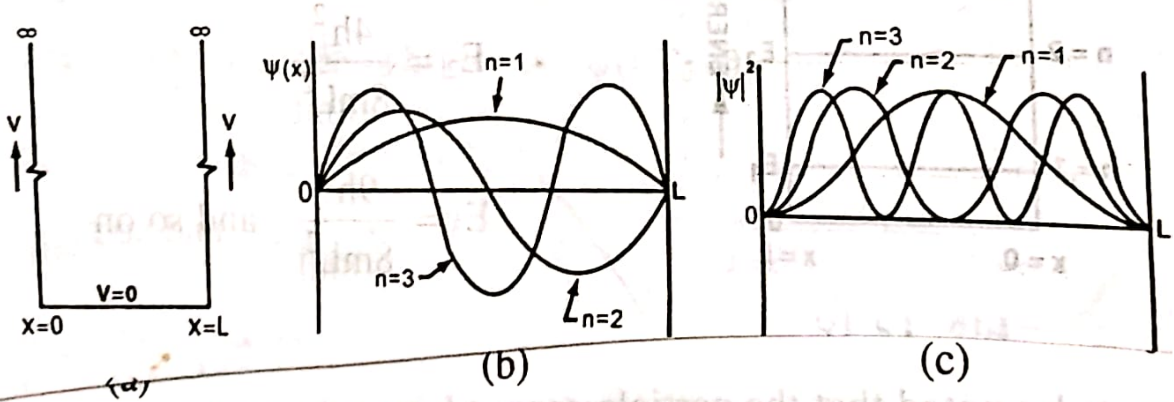


Fig. 15.17

***Example 15.12** Calculate the permitted energy levels of an electron, in a box 1 A.U. wide.*

Soln.

The permitted electron energies are given by

$$E_n = \frac{n^2h^2}{8mL^2} \qquad \text{Here} \quad m = 9.1 \times 10^{-31}\,kg$$

$$L = 1\text{ A.U.} = 10^{-10}\,m$$

$$= \frac{n^2(6.62\times10^{-34})^2}{(8)(9.1\times10^{-31})(10^{-10})^2}$$

$$= 6\times10^{-18}\, n^2 J = 38\, n^2\, eV.$$

The minimum energy, the electron can have, corresponds to n=1. Or

$$E_1 = 38\text{ eV}.$$

The other values of energy are

$$E_2 = 4E_1 = 152\text{ eV};\; E_3 = 9E_1 = 342\text{ eV and so on.}$$

***Example 15.13** Consider a 1-μg speck of dust moving back and forth between two rigid walls separated by 0.1mm. It moves so slowly that it takes 100s for the particle to cross this gap. What quantum number describes this motion?*

Soln.

The energy of the particle is

$$E\,(=K) = \frac{1}{2}mv^2 \qquad \text{Here } m = 1\mu g = 1\times10^{-9}\,kg$$

$$= \frac{1}{2}(1\times10^{-9}\,kg)(1\times10^{-6}\,m/s)^2 \qquad v = \frac{0.1mm}{100s} = \frac{1\times10^{-4}\,m}{100s}$$

$$= 5\times10^{-22}\,J. \qquad = 10^{-6}\,m/s.$$

From the relation $E = \frac{n^2h^2}{8mL^2}$, we find the quantum number corresponding to this energy to be

$$= \frac{[illegible]}{h}\sqrt{8mE} = \frac{[illegible]10^{-4}}{(6.63\times10^{-34}\,J.s)}\sqrt{(8)(10^{-9}\,kg)(5\times10^{-22}\,J)}$$

$$\cong 3\times10^{14}$$

It is a very large number. Experimentally it is impossible to distinguish between $n = 3\times10^{14}$ and $n = 3\times10^{14}+1$, so that the quantized nature of the motion would never reveal itself. Although the mass and the kinetic energy of the speck of dust are both extremely small, it is still a gross macroscopic object when compared to an electron.

Example 15.14 *An electron is trapped in a one-dimensional region of length 1.0×10^{-10}m (a typical atomic diameter). (a) How much energy must be supplied to excite the electron from the ground state to the first excited state? (b) In the ground state, what is the probability of finding the electron in the region from $x = 0.090\times10^{-10}$ m to 0.110×10^{-10} m? (c) In the first excited state, what is the probability of finding the electron from $x = 0$ and $x = 0.250\times10^{-10}$ m?*

Soln.

(a) $E_1 = \dfrac{h^2n^2}{8mL^2}$ Here $n = 1$

$m = 9.1\times10^{-31}$ kg

$L = 1.0\times10^{-10}$ m

$$= \frac{(6.626\times10^{-34}\,\text{Js})^2}{(8)(9.1\times10^{-31}\,\text{kg})(1.0\times10^{-10}\,\text{m})^2}$$

$$= 6.0\times10^{-18}\,\text{J} = 37\text{ eV}$$

Now $E_2 = (2)^2E_1 = 4E_1$

$\therefore \quad \Delta E = 4E_1 - E_1 = 3E_1 = (3\times37)\text{eV} = 111\text{ eV}.$

(b) Now probability is given by

$$\int_{x_1}^{x_2}\psi^2dx = \frac{2}{L}\int_{x_1}^{x_2}\sin^2\frac{\pi x}{L}dx$$

$$= \left[\frac{x}{L} - \frac{1}{2\pi}\sin\frac{2\pi x}{L}\right]_{x_1}^{x_2}$$

$$= \left[\frac{x}{L} - \frac{1}{2\pi}\sin\frac{2\pi x}{L}\right]_{0.090\times10^{-10}\text{m}}^{0.110\times10^{-10}\text{m}}$$

$$= 0.0038 = 0.38\%$$

(c) $\text{Probability} = \int_{x_1}^{x_2} \left(\frac{2}{L}\right)\sin^2\frac{2\pi x}{L}\,dx$ $\qquad n = 2$

$$= \left[\frac{x}{L} - \frac{1}{4\pi}\frac{4\pi x}{L}\right]_{x_1}^{x_2} \qquad x_2 = 0.250\times10^{-10}\text{m}$$

$$= 0.25$$

***Example 15.15** A particle is moving in a one-dimensional box (of infinite height) of width 10 Å. Calculate the probability of finding the particle within an interval of 1 Å at the centre of the box, when it is in its state of least energy.*

Soln.

The wave function of the particle in the ground state (n=1), is

$$\psi = \sqrt{\frac{2}{L}}\sin\frac{\pi x}{L}$$

Probability of finding the particle in unit interval at the centre of the box (x = L/2) is

$$P = \psi_1^2 = \left[\sqrt{\frac{2}{L}}\sin\frac{\pi(L/2)}{L}\right]^2$$

$$= \frac{2}{L}\sin^2\frac{\pi}{2} = \frac{2}{L}$$

Probability of finding the particle within an interval of Δx at the centre of the box

$$= |\psi_1|^2\Delta x = \frac{2}{L}.\Delta x \qquad \text{Here } L = 10\text{ Å} = 10\times10^{-10}\text{m}$$

$$= \frac{2}{10\times10^{-10}}\times10^{-10} = 0.2 \qquad \Delta x = 1\text{ Å} = 10^{-10}\text{m}$$

Linear Harmonic Oscillator

Another situation that can be handled easily using the Schrödinger equation is the one-dimensional simple harmonic oscillator. The classical oscillator is an object of mass m attached to a spring of force constant k. If x is the displacement of the object from its equilibrium position (which we take to be x = 0), then a restoring force F = – kx acts on the object. By using Newton's laws, it can be shown that the frequency and time period of the oscillator are $w = \sqrt{k/m}$ and $T = 2\pi\sqrt{k/m}$ respectively. The oscillator has its maximum kinetic energy at x = 0, the kinetic energy vanishes at the turning points $x = \pm A_0$, where A_0 is the amplitude of the motion. At the turning points the oscillator comes to rest for an instant and then reverses its direction of motion. Thus, the motion remains confined to the region $-A_0 \le x \le +A_0$.

The potential energy associated with a force F = – kx is given by $U = \frac{1}{2}kx^2$. The Schrödinger equation for the harmonic oscillator, therefore, becomes

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}kx^2\psi = E\psi \qquad \text{(i)}$$

Since we are working in one dimension, U and ψ are functions only of x. It is difficult to solve eqn.(i) directly. The solution must approach zero as $x \to \pm\infty$, and in the limit $x \to \pm\infty$ the solutions of eqn.(i) behave like exponentials of $-x^2$. We therefore try $\psi(x) = Ae^{-ax^2}$, where A and a are constants that are determined by evaluating eqn(i) for this choice of ψ(x).

Let us begin by evaluating $\frac{d^2\psi}{dx^2}$

$$\frac{d\psi}{dx} = -2ax(Ae^{-ax^2})$$

$$\frac{d^2\psi}{dx^2} = -2a(Ae^{-ax^2}) - (2ax)(-2ax)(Ae^{-ax^2})$$

$$= -2a(Ae^{-ax^2}) + 4a^2x^2(Ae^{-ax^2})$$

Substituting the values of $\psi(x)$ and $\frac{d^2\psi}{dx^2}$ in eqn.(i), we get

$$-\frac{\hbar^2}{2m}(-2aAe^{-ax^2} + 4a^2x^2Ae^{-ax^2}) + \frac{1}{2}kx^2(Ae^{-ax^2})$$

$$= E.\ Ae^{-ax^2} \qquad \text{(ii)}$$

Canceling Ae^{-ax^2} on both sides of eqn(ii), we obtain

$$\frac{\hbar^2 a}{m} - \frac{2a\hbar^2}{m}x^2 + \frac{1}{2}kx^2 = E \qquad \text{(iii)}$$

Eqn (iii) is *not* an equation to be solved for x, because we are looking for a solution which is valid for *any* x, not just for one specific value. For the solution of eqn (iii) to hold for any x, the coefficients of x^2 must cancel and the remaining constants must be equal. Thus

$$-\frac{2a^2\hbar^2}{m} + \frac{1}{2}k = 0 \qquad \text{(iv)}$$

and $$\frac{\hbar^2 a}{m} - E = 0 \qquad \text{(v)}$$

which yield

$$a = \frac{\sqrt{km}}{2\hbar} \qquad \text{(vi)}$$

and $$E = \frac{1}{2}\hbar^2\sqrt{k/m} \qquad \text{(vii)}$$

The expression for energy can also be written in terms of the classical frequency $w_0 = \sqrt{k/m}$ as

$$E = \frac{1}{2}\hbar w_0 \qquad \text{(viii)}$$

Using normalization condition, it can be shown that the constant

$$A = \left(\frac{mw_0}{\hbar\pi}\right)^{1/4}$$

The solution just obtained is illustrated in Fig.15.18. The striking feature of the solution is that there is a non-zero probability of finding

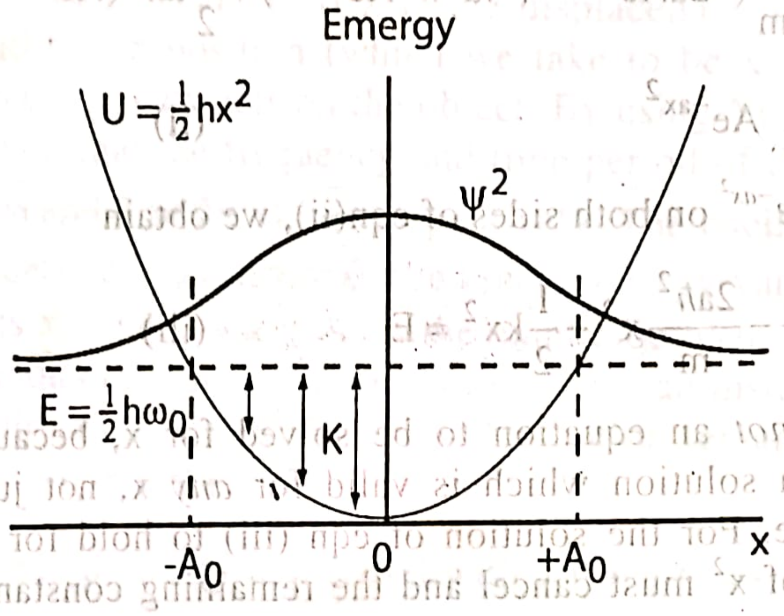


Fig. 15.18

the particle beyond the classical turning points $x = \pm A_0$. The total energy E is constant, and beyond $x = \pm A_0$, the potential energy is greater than E, so that the kinetic energy would become *negative.* This is impossible in classical physics, and so the classical *particle* can never be found at $|x| > A_0$. It is, however, possible for the quantum *wave* to penetrate into the classically forbidden region.

The solution that we have found corresponds to the *ground state* of the oscillator. The mathematically difficult general solution is of the form $\psi_n(x) = Af_n(x)e^{-ax^2}$, where $f_n(x)$ is a polynomial in which the highest power of x is x^n.

The corresponding energies are

$$E_n = (n + \frac{1}{2})\hbar w_0 \qquad (n = 0, 1, 2, \ldots) \qquad (15.19)$$

It may be noted that these levels, illustrated along with their probability densities in Fig. 15.19, are *uniformly spaced,* in contrast to the one-dimensional particle in a box. All of the solutions have the property of penetration of probability density into the forbidden

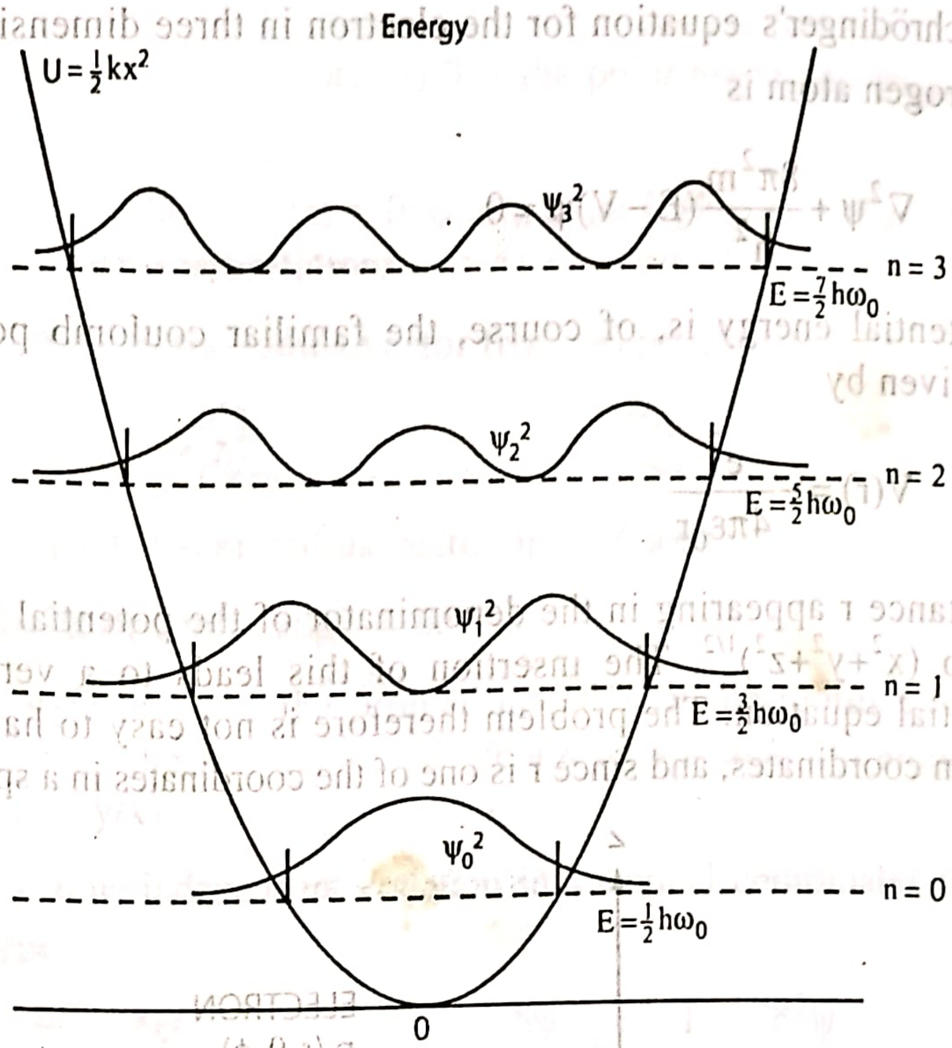


Fig. 15.19

region beyond the classical turning points. The probability density oscillates, somewhat like a sine wave, between the turning points, and decreases like e^{-2ax^2} to zero beyond the turning points.

The Hydrogen Atom

The hydrogen atom consists of a proton around which the electron revolves. The proton is assumed to be at rest at the origin of a rectangular coordinate system while the electron is orbiting around the proton under the influence of the attractive coulomb field of the system, the radius of the orbit being r.

From the point of view of quantum mechanics, the electron is represented by a wave system bounded by the potential well of the coulomb field. This circumstance results in a set of permitted standing wave systems, each corresponding to a particular possible value of the total energy.

Schrödinger's equation for the electron in three dimensions for the hydrogen atom is

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0$$

The potential energy is, of course, the familiar coulomb potential and is given by

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$$

The distance r appearing in the denominator of the potential term is equal to $(x^2+y^2+z^2)^{1/2}$. The insertion of this leads to a very ugly differential equation. The problem therefore is not easy to handle in Cartesian coordinates, and since r is one of the coordinates in a spherical

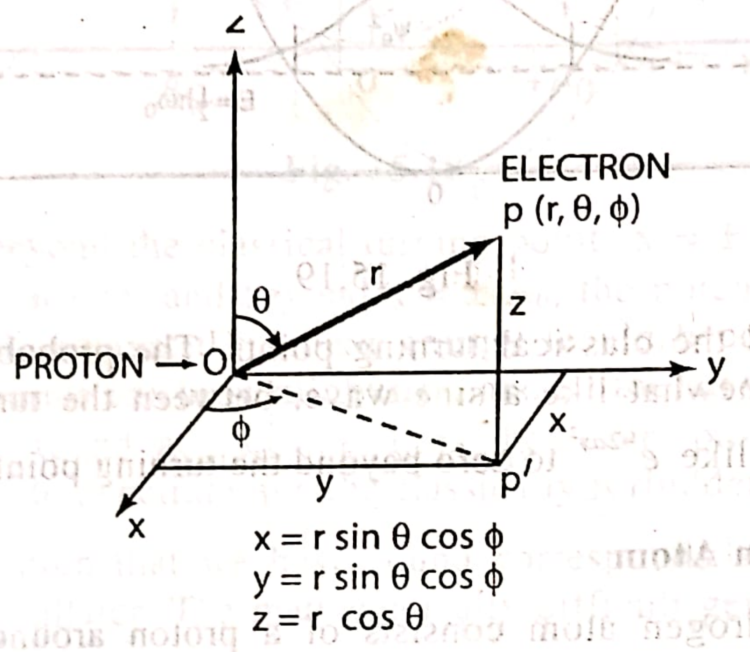


Fig. 15.20

coordinate system, it will be more convenient to use the spherical polar coordinate system. The relationship between the spherical coordinates r, θ and φ and the Cartesian coordinates x,y,z can be read directly from Fig. 15.20. They are

$$x = r\sin\theta\cos\phi$$

$$y = r\sin\theta\sin\phi$$

$$z = r\cos\theta$$

where (i) r is the radius vector, (ii) θ is the polar angle and (iii) ϕ is the azimuth angle.

The spherical polar coordinates r, θ, ϕ of a point P shown in Fig. 15.20 have the following interpretations:

r = length of the radius vector from origin O to point P

$= \sqrt{x^2 + y^2 + z^2}$

θ = angle between radius vector and +Z-axis

= polar or zenith angle = $\cos^{-1} z/r$

ϕ = angle between the projection (OP′) of the radius vector OP on the XY-plane and the +X axis = azimuth angle = $\tan^{-1}(y/x)$

Schrödinger's time-independent equation in spherical coordinates then takes the form

$$\frac{1}{r^2}\frac{\delta}{\delta r}\left(r^2\frac{\delta\psi}{\delta r}\right)+\frac{1}{r^2\sin\theta}\frac{\delta}{\delta\theta}\left(\sin\theta\frac{\delta\psi}{\delta\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\delta^2\psi}{\delta\varphi^2}$$

$$+\frac{8\pi^2 m}{h^2}(E-V)\psi = 0 \qquad \text{(i)}$$

Substituting $V = -e^2/4\pi\varepsilon_0 r$ and multiplying all throughout by $r^2\sin^2\theta$, we obtain

$$\sin^2\theta\frac{\delta}{\delta r}\left(r^2\frac{\delta\psi}{\delta r}\right)+\sin\theta\frac{\delta}{\delta\theta}\left(\sin\theta\frac{\delta\psi}{\delta\theta}\right)+\frac{\delta^2\psi}{\delta\phi^2}$$

$$+\frac{8\pi^2 mr^2\sin^2\theta}{h^2}\left(E+\frac{e^2}{4\pi\varepsilon_0 r}\right)\psi = 0 \qquad \text{(ii)}$$

Eqn. (ii) is the partial differential equation for the wave function ψ of the electron in a hydrogen atom.

Separation of variables

The wave function now depends on r, θ and φ. Here we look for solutions in which the wave function $\psi(r,\theta,\phi)$ has the form of a product of three different functions.

$\psi_1(r)$ is a function which depends on r alone

$\psi_2(\theta)$ is a function which depends on θ alone.

$\psi_3(\phi)$ is a function depending on ϕ alone.

We, therefore, assume that

$$\psi(r,\theta,\phi) = \psi_1(r)\psi_2(\theta)\psi_3(\phi) \qquad \text{(iii)}$$

Substituting this value of ψ in (ii), dividing throughout by $\psi_1\,\psi_2\,\psi_3$ and rearranging we get

$$\frac{\sin^2\theta}{\psi_1}\frac{d}{dr}\left(r^2.\frac{d\psi_1}{dr}\right)+\frac{\sin\theta}{\psi_2}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_2}{d\theta}\right)+\frac{8\pi^2mr^2\sin^2\theta}{h^2}\left(E+\frac{e^2}{4\pi\varepsilon_0 r}\right)$$

$$=-\frac{1}{\psi_3}\frac{d^2\psi_3}{d\phi^2} \qquad \text{(iv)}$$

Left hand side of eqn. (iv) is a function of r and θ alone while the right hand side is a function of ϕ alone. The only way in which this equation can be satisfied for all values of r,θ and ϕ is to have both sides of the equation equal to the same constant. It is convenient to call this constant m_l^2. Therefore the differential equation for the function ψ_3 is

$$-\frac{1}{\psi_3}\frac{d^2\psi_3}{d\phi^2}=m_l^2$$

This gives the **azimuthal equation**

$$\frac{d^2\psi_3}{d\phi^2}+m_l^2\psi_3=0 \qquad (15.20)$$

When we substitute m_l^2 for the right hand side of eqn.(iv), divide the entire equation by $\sin^2\theta$ and rearrange the various terms, we obtain

$$\frac{1}{\psi_1}\frac{d}{dr}\left[r^2\frac{d\psi_1}{dr}\right]+\frac{8\pi^2mr^2}{h^2}\left[E+\frac{e^2}{4\pi\varepsilon_0 r}\right]$$

$$= \frac{m_l^2}{\sin^2\theta} - \frac{1}{\psi_2 \sin\theta} \frac{d}{d\theta}\left[\sin\theta \frac{d\psi_2}{d\theta}\right]$$

Again we have an equation in which different variables appear on each side. This again requires that both sides should be equal to the same constant. Let us call this constant $l\ (l+1)$. Therefore, the equations for the functions ψ_2 and ψ_1 are,

$$\frac{m_l^2}{\sin^2\theta} - \frac{1}{\psi_2 \sin\theta} \frac{d}{d\theta}(\sin\theta \frac{d\psi_2}{d\theta}) = l\ (l+1) \tag{15.21}$$

Eqn.(15.21) is called the **polar equation.**

$$\frac{1}{\psi_1} \frac{d}{dr}(r^2 \frac{d\psi_1}{dr}) + \frac{8\pi^2 mr^2}{h^2}(E + \frac{e^2}{4\pi\varepsilon_0 r}) = l\ (l+1) \tag{15.22}$$

Eqn.(15.22) is called the **radial equation.**

Thus, Schrödinger wave equation for hydrogen atom has now been separated into three equations [(15.20), (15.21) and (15.22)] each dependent on only one of the coordinates.

The Azimuthal equation. The solution of the azimuthal eqn. [(v)] is given by

$$\psi_3(\phi) = Ae^{im_l\phi} \tag{15.23}$$

For this function to be single valued, its value must be the same for $\phi=\phi$ and $\phi=\phi+2\pi$

$$\text{or} \quad Ae^{im_l\phi} = Ae^{im_l(\phi+2\pi)}$$

To satisfy this condition, m_l must be zero or a positive or negative integer. Therefore, the permitted values of m_l are

$$m_l = 0, \pm1, \pm2, \pm3, \tag{15.24}$$

The constant m_l is known as the **magnetic quantum number** of the hydrogen atom.

The polar equation

The differential equation (15.21) has a rather complicated solution in terms of polynomials called the **associated Legendre polynomials** represented by $P_l\ m_l\ (\cos\theta)$. For these polynomials to exist, the condition that must be satisfied is that l must be an integer equal to or greater than $[m_l]$. This condition may be expressed as follows:

$$m_l = 0, \pm1, \pm2, \pm3, \ldots \pm l \qquad (15.25)$$

The constant l is known as the **orbital quantum number.**

The radial equation. The solutions of the radial equation (15.22) are given by the standard mathematical functions known as **associated Laguerre polynomials** $L_{n,l}(r)$. Eqn.(15.22) can be solved only when E is positive or has one of the negative values E_n specified by

$$E_n = -\frac{me^4}{8\varepsilon_0^2h^2}\left(\frac{1}{n^2}\right) = (-13.6)\frac{1}{n^2}\text{eV} \qquad (15.26)$$

The negative values E_n signify that the electron is bound to the atom.

Eqn. (15.26) is in complete agreement with the predictions of Bohr's atom model. Here n is again an integer whose value must be greater than or equal to $(l+1)$ i.e., $n \geq (l+1)$. Thus we have arrived at a system of three quantum number for the simple model of hydrogen atom. The quantum numbers are related in the following manner:

total quantum number	$n = 1,2,3,\ldots$
orbital quantum number	$l = 0,1,2,\ldots (n-1)$
magnetic quantum number	$m_l = 0, \pm1, \pm2, \pm3, \ldots \pm l.$

EXERCISE

1. What led de Broglie to propose his matter-wave hypothesis? Discuss briefly the wave nature of matter and obtain an expression for the de Broglie wavelength of matter waves.

2. Describe Davisson and Germer experiment for the study of electron diffraction. Discuss the significance of the results obtained.

3. Describe the experiment of G.P. Thomson on the diffraction of electron and discuss the results obtained.

4. What is meant by matter waves? What are the characteristics of these waves? Describe an experiment which support the concept of matter waves.

5. Describe an experiment to confirm the existence of de Broglie waves. Using the concept of de Broglie waves, establish Bohr's quantum condition of stationary waves.

6. State and explain the uncertainty principle and illustrate it by means of a suitable experiment.

7. Discuss the significance and the utility of Heisenberg's uncertainty principle.

8. State the different forms of the uncertainty relations. Outline an idealized experiment to bring out its significance. Can an electron stay inside a nucleus? Explain.

9. Describe the principle, construction and working of an electron microscope. Compare this instrument with an optical microscope.

10. What are the postulates of wave mechanics? Derive Schrödinger's equation. Discuss the significance of the wave function.

11. Write down Schrödinger equation and solve it to calculate the values of the energy of a particle in a one-dimensional box. Indicate graphically the first three wave functions for such a particle.

12. Explain the requirements that are imposed on a physically acceptable wave functions. What do you understand by the terms "eigen value" and "eigen function"?

13. Write down the Schrödinger equation for a particle in one-dimensional potential well described by

    $V(x) = 0$ for $-a<x<a$

    $V(x) = \infty$ elsewhere.

    Find the allowed energies and the wave functions. Why is the ground state energy not zero?

14. Solve the Schrödinger wave equation for a particle restricted to travel along the x-axis within a box with impenetrable wall, and determine its energy eigen value and normalized wave function.

15. Deduce the Schrödinger wave equation for a free particle moving in a field of constant potential V_0. Discuss the significance of wave function.

16. What are the postulates of wave mechanics? Give the physical interpretation of wave-function.

17. Establish Schrödinger equation for a linear harmonic oscillator and slove it to obtain its eigen values and eigen functions. In what way is the quantum mechanical description of a simple harmonic oscillator different from classical description?

18. Write the Schrödinger wave equation of the hydrogen atom in polar coordinates. Explain the origin and significance of the quantum numbers n, l and m_l.

19. Solve the Schrödinger equation for hydrogen atom. Neglect electron spin and assume the nucleus to be stationary.

20. Find the de Broglie wavelength of electrons of kinetic energy 1eV. **[122.2A^0]**

21. Determine the wavelength associated with an electron having K.E. equal to 1 MeV. **[0.00726A^0]**

22. A proton and an α particle have the same K.E. (i) How do their speeds compare? (ii) How do their momenta compare? (iii) How do their de Broglie wavelengths compare? **[(i) $v_p=2v_\alpha$; (ii) $p_p=\frac{1}{2}p_\alpha$; (iii) $\lambda_p=2\lambda_\alpha$]**

23. A beam of monoenergetic neutrons corresponding to 27°C is allowed to fall on a crystal. A first order reflection is observed at a glancing angle of 30°. Calculate the interplanoar spacing of the crystal. Given: $h=6.62\times 10^{-34}$ Js; $m_n=1.67\times 10^{-27}$kg, Boltzman's constant $k = 1.38\times10^{-23}$ joule/degree. **[$d=1.82\times 10^{-10}$m]**

24. A certain excited atom of hydrogen is known to have a life time of 2.5×10^{-14}sec. What is the minimum error with which the energy of the excited state can be measured? If the centre of a hydrogen atom be located with a precision of 0.01A.U., what is the corresponding uncertainty in its velocity? **[4.2×10^{-21}joule; 6.3×10^4m/sec]**

25. The position and momentum of a 1 KeV electron are simultaneously determined. If its position is located to within 1A.U., what is the percentage of uncertainty in its momentum? **[6.173%]**

26. Compare the uncertainties in the velocities of an electron and a proton confined in a 10 A.U. box. **[$1.16\times10^5\ ms^{-1}$; $63\ ms^{-1}$]**

27. Applying uncertainty principle, show that the presence of protons in a nucleus is entirely plausible.

28. An electron is confined to move between two rigid walls separated by 10^{-9}m. Find the de Broglie wavelengths representing the first three allowed energy states of the electrons and the corresponding energies. **[(i) 20A°, 10 A°, 6.7 A° (ii) 0.38 eV, 1.52 eV, 3.42 eV]**

29. A particle is moving in a one-dimensional potential box of infinite height. What is the probability of finding the particle in a small interval Δx at the centre of the box when it is in the energy state, next to the least energy state? **[Δx=0].**

CHAPTER XVI

RADIOACTIVITY

Radioactivity is a spontaneous and self-disruptive activity exhibited almost entirely by some heavy elements of atomic weights greater than about 206, occurring in nature. The elements that exhibit this property are called *radioactive elements.* The activity consists in the emission of powerful radiations composed of three distinct kinds of radiation, known as *alpha* (α), *beta* (β) and *gamma* (γ) rays. The result of this activity is the breaking up of the element for good, *i.e.* an irreversible self-disintegration. The activity is spontaneous in the sense that it arises solely from intrinsic natural causes, unaffected by any external agent, physical or chemical. As a matter of fact, heat enough to melt any element, cold enough to freeze any substance, strong electric fields to tear electrons out, intense magnetic fields, all these have been tried on radioactive elements. But the activity remains unaltered in every instance. Further, the activity is not instantaneous but prolonged over a certain period of time; otherwise it could not have been discovered at all.

16.1 Natural and Artificial Radioactivity

With the advent of modern technique of artificial transmutation of elements, it has become possible to induce radioactivity in many other elements much lighter that those occurring in nature. This has necessitated a distinction between *natural activity i.e.*, radioactivity exhibited by elements as found in nature, and *induced* or *artificial radioactivity i.e.*, radioactivity produced artificially. But whatever be its origin, natural or artificial, the activity is always spontaneous. Usually, but not always, the artificial radioactive elements have short half-lives. They emit electrons, positrons, and other particles, as well as gamma rays in their disintegrations.

16.2 Discovery of radioactivity

The discovery of radioactivity was quite accidental. Henri Becquerel, a French physicist, was investigating an interesting point in connection with X-rays, *viz.* the possible relationship between the fluórescence of the glass walls of an X-ray tube accompanying emission of X-rays and the phosphorescence of certain fluorescent salts activated by sunlight. This was April, 1896, only a few months after the discovery

of X-rays by Roentgen (November, 1895). Becquerel was actually experimenting with the double sulphate of uranium and potassium. One day, as no sunlight was available due to cloudy weather, he left the substance wrapped up in a paper on some photographic plates also covered in black paper inside a drawer. A few days later, when the photographic plates were developed, the plate showed images of uranium crystals, although the plates were well packed in black paper and the salt itself had not been exposed to sunlight. Becquerel immediately surmised that the uranium salt *even in the* dark might be emitting radiations powerful enough to penetrate the paper wrappers and affect the photographic plates. To see whether his guess was correct, he performed a series of experiments by placing the uranium salt and the photographic plates inside a light tight box and arranging them either in contact or separating them by thin sheets of aluminum. The results clearly showed that the active radiations came from the uranium salt itself and that exposure to sunlight had no influence whatever on the phenomenon. He was able to prove further that the radiations were emitted by uranium, whether it was in solid form or in solutions, irrespective of its state of chemical combination and physical conditions of temperatures. The name *Becquerel rays* was first given to the radiations emitted by uranium. Subsequently it was changed to *radioactivity* to designate all such radiations emitted not only by uranium but by several other substances which were soon discovered such as radium, thorium, polonium and actinium.

The most significant discovery was radium. In 1898 Mme. Curie and her husband, Pierre Curie found that pitchblende, a uranium mineral, was many times more active than a pure uranium salt. By fractional crystallisation, the Curies succeeded in separating from the mineral two new highly active elements named *radium* and *polonium*. Polonium was many times more active than uranium but the activity of radium was enormously greater – more than a million times that of uranium for equal weights. The process of separation of these radioactive substances from the original mineral was an extremely difficult task, a ton of uranium ore, for instance, yielding only a few decigrams of radium. In the same year, 1898, Schmidt discovered *thorium*. Debierne in 1900 separated *actinium* from pitchblende. *Radiothorium* and *mesothorium* were discovered by Hahn in 1905 and 1907 respectively. Hahn and Meitner discovered *protoactinium* in 1918. Many more discoveries of radioactive substance followed.

16.3 Alpha, Beta and Gamma Rays

Soon after the discovery of radioactivity by Becquerel, Lord Rutherford and his collaborators at Cambridge investigated the nature of the radiations emitted by radioactive substance by studying the penetrating power of the radiations emitted by uranium. They used sheets of tin foil as absorber and gold leaf electroscope to measure the intensity of ionization produced by the transmitted radiations. They showed that the original radiations emitted by uranium were of two distinct kinds. One is very soft, easily absorbed and capable of producing intense ionization. Rutherford called it α-rays. The other is much more penetrating than the first but much less effective in producing ionisation. It was called β-rays. Subsequently Villard discovered that there is also a third kind of radiation much more penetrating than α-rays or β-rays. It was named γ-rays.

The existence of these three distinct types of radiations was demonstrated by a simple experiment first devised by Mme. Curie.

A small quantity of a radioactive element, say radium (R), is placed at the bottom of a small hole drilled in a lead block [Fig. 16.1]. This helps in obtaining a narrow and fairly parallel beam of radiation which emerge from the mouth of the hole. Other rays entering the walls of the lead block are absorbed before reaching the surface. A photographic plate (P) is placed a short distance above the lead block to receive the radiations. A strong magnetic field is applied at right angles to the plane of the figure and directed away form the reader.

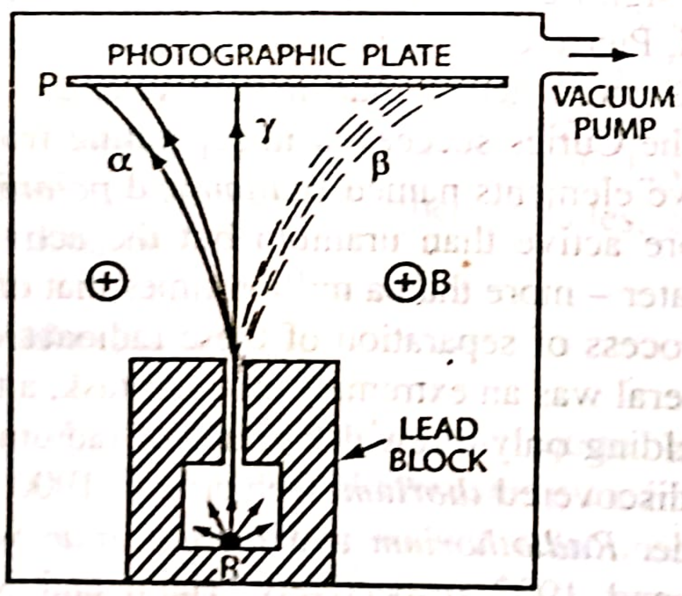


Fig. 16.1

When the photographic plate is developed after a fairly long exposure, three blackenings or traces are found on it. The one lying to the left is produced by the positively charged α-particles while the one to the right is produced by the negatively charged β-particles. The γ-rays are not deflected and hit the plate P straight. This proves that γ-rays are uncharged or neutral rays.

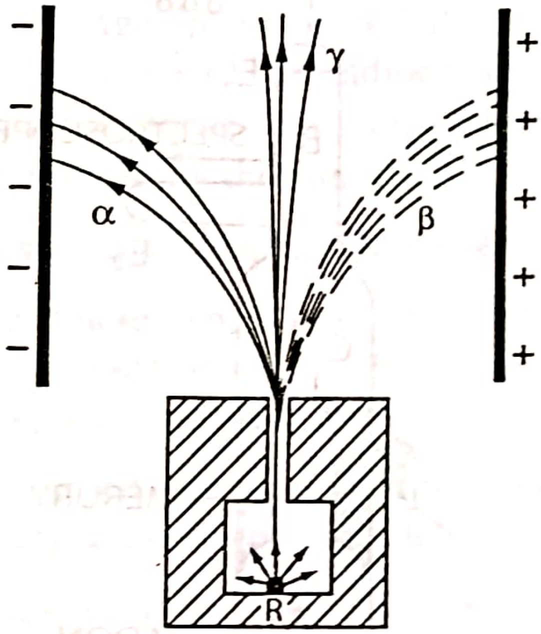


Fig. 16.2

If instead of the magnetic field, an electrostatic field is applied, similar results will be obtained [Fig. 16.2]. The negatively charged β-particles are deflected towards the positively charged plate whereas the positively charged α-particles are deflected towards the negative plate. As expected the γ-rays do not bend at all.

Properties of α-particles

(1) α-particles are positively charged particles and their charge and mass have been determined. The ratio of charge to mass (e/m) of α-particles was found by Rutherford to be 4.823×10^7 C/kg. The charge on each α-particle was found to be twice that of a hydrogen ion (3.202×10^{-19} C). The mass of the α-particle was shown to be four times that of hydrogen. Thus the α-particle was identified as the nucleus of the helium atom (ionised helium atom).

Experimental proof:

A thin-walled glass tube A is filled with radon and sealed into an outer tube B which is highly evacuated (Fig. 16.3). The tube A is so thin that α-particles can pass through it. α- particles are collected in the tube C. The mercury level in the tube C is raised by connecting

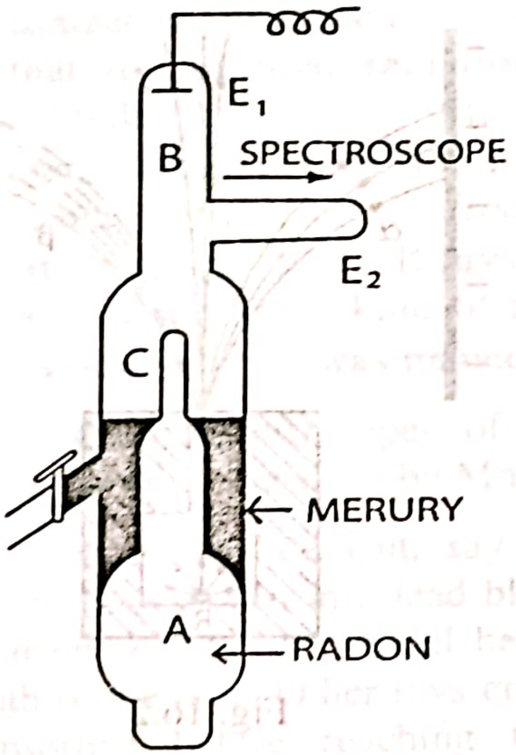


Fig. 16.3

it to a mercury reservoir. In this way, α- particles collected in the tube C are pushed into the tube B. The electrodes E_1 and E_2 are connected to the secondary of an induction coil. A discharge was passed between the electrodes. The light coming out of the tube was examined with a spectroscope which clearly showed the spectral lines of helium. Control experiments showed that ordinary helium gas could not penetrate through the thin walls of the tube A. This spectroscopic evidence, therefore, proves conclusively that α-particles are nuclei of helium.

(2) α- particles are ejected from radioactive nuclei with high velocities, ranging from 1.42×10^7 to 2.05×10^7 m/s. They move along straight lines and produce intense ionization in the gases through which they pass. Their ionization power is 100 times greater than that of β-rays and 10,000 times

greater than that of γ-rays. Hence, they produce dense cloud chamber tracks that resemble the vapour trails of a high flying aircraft.

(3) Out of the three radioactive radiations, they have the least penetrating power. α- particles are easily absorbed by thin metal sheets and by few centimeters of a gas at atmospheric pressure. The range in air at N.T.P. is about 5 cm; but it depends on the nature of the radioactive elements. Its value varies from 3.4 cm for α- particles emitted by radium to 8.6 cm for those emitted by thorium C′.

(4) They affect a photographic plate but the effect is very feeble.

(5) They produce fluorescence when they fall on substance like barium platinocyanide or zinc sulphide. The fluorescence can be observed through a spinthariscope *i.e.*, a low power microscope.

(6) They are deflected by electric and magnetic fields. This proves that they are charged particles.

(7) α- particles are scattered while passing through thin sheets of mica, aluminium and gold foil etc. Generally, the deflection is not more than 2° to 3°; but sometimes deflection of more than 90° are produced by heavy elements like gold.

Properties of β-rays

1. β-rays consists of ordinary electron – they possess the same negative charge and mass as that of an electron but they are of nuclear origin.

2. β-particles do not emerge out of the nucleus with the same range and velocity; rather they have a continuous range of velocities from 0.3c to 0.99c. The ratio of charge to mass (e/m) of electron is found to decrease at high velocities indicating an increase of mass of electron with increase of its velocity.

3. They produce ionization in air; but their ionization power is $1/100^{th}$ of that of α-particles. Thus the range of β-particles is large.

4. They affect photographic plates more strongly than α-particles.
5. They produce fluorescence in barium platinocyanide, calcium tungstate, willemite, etc.
6. They are deflected by electric and magnetic fields. Their direction of deflection indicates that they are negatively charged particles.
7. They penetrate through thin metal foils and their penetrating power is greater than that of α-rays.

Properties of γ-rays

Some of the important properties of gamma rays are mentioned below:

(i) Gamma rays are electromagnetic radiations of very short wavelength ranging from 0.5 Å to 0.005 Å. They are not charged particles and they travel with the velocity of light.

(ii) They produce fluorescence and affect photographic plates more intensely than β-rays.

(iii) They ionize the gas through which they pass but the ionization produced is very small.

(iv) They are not affected by electric and magnetic fields.

(v) Like X-rays, they can be diffracted by crystals.

(vi) They are more penetrating than even β-rays. They can even pass through an iron plate of 30 cm thickness. While passing through matter, they are absorbed according to an exponential law $I = I_0e^{-\mu x}$ where x is the thickness of the matter and μ is the absorption coefficient. There are three processes by which γ-rays are absorbed by matter.

(a) **Photoelectric effect:** During a photoelectric interaction, the total energy of the γ-ray photon is transferred to an inner electron of an atom (Fig. 16.4). The electron is ejected from the atom with kinetic energy E_k, where E_k equals the photon energy hv minus the binding energy E_B required to remove the

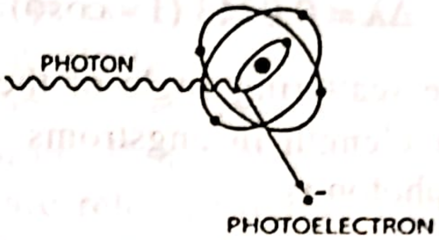


Fig. 16.4

electron from the atom.

$$E_k = h\nu - E_B$$

The ejected electron is called a **photoelectron**. When an electron is ejected from K-shell, then the vacancy so created may be filled by an electron from the L-shell or M-shell. When the electron jumps from either the L-shell or M-shell to the K-shell, energy is radiated out in the form of X-rays which are found to be identical with those emitted by the same element in an X-ray tube.

(b) Compton scattering: Gamma rays having medium photon energy (30 KeV to 30 MeV) usually interact with matter by a process known as Compton scattering. In Compton interaction, a part of the energy of an incident photon is transferred to a loosely bound or "free" electron within the attenuating medium. The electron recoils at an angle θ with respect to the motion of the incident photon, and the photon is scattered at an angel ϕ (Fig. 16.5). The kinetic energy of the

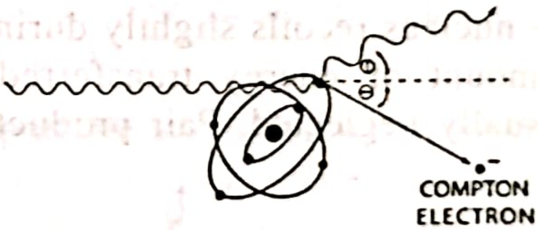


Fig. 16.5

recoil or Compton electron equals the energy lost by the photon, assuming that the binding energy of the electron is negligible. Both θ and ϕ decrease with increasing energy of the incident photon.

During a Compton interaction, the change in wavelength of the γ-ray photon is

$$\Delta\lambda = 0.0243\,(1 - \cos\phi)$$

where ϕ is the scattering angle of the photon and $\Delta\lambda$ is the change in wavelength in angstroms. The wavelength λ' of the scattered photon is

$$\lambda' = \lambda + \Delta\lambda$$

where λ is the wavelength of the incident photon. The energies $h\nu$ and $h\nu'$ of the incident and scattered photons are

$$h\nu\,(\text{keV}) = \frac{12.4}{\lambda(\text{A}^0)}$$

$$\text{and} \quad h\nu'(\text{keV}) = \frac{12.4}{\lambda'(\text{A}^0)}$$

(c) Pair production: In this process of absorption, a γ-ray photon disappears in the Coulomb field of an atomic nucleus and an electron-positron pair appears in its place. Since the energy equivalent to the mass of an electron is 0.51 MeV, the creation of two electrons requires 1.02 MeV. Consequently, photons with energy less than 1.02 MeV do not interact by pair production. During pair production, energy in excess of 1.02 MeV is released as kinetic energy of the two electrons.

$$h\nu\,(\text{MeV}) = 1.02 + (E_k)_{e^-} + (E_k)_{e^+}$$

Although the nucleus recoils slightly during pair production, the small amount of energy transferred to the recoiling nucleus is usually neglected. Pair production is depicted in Fig. 16.6.

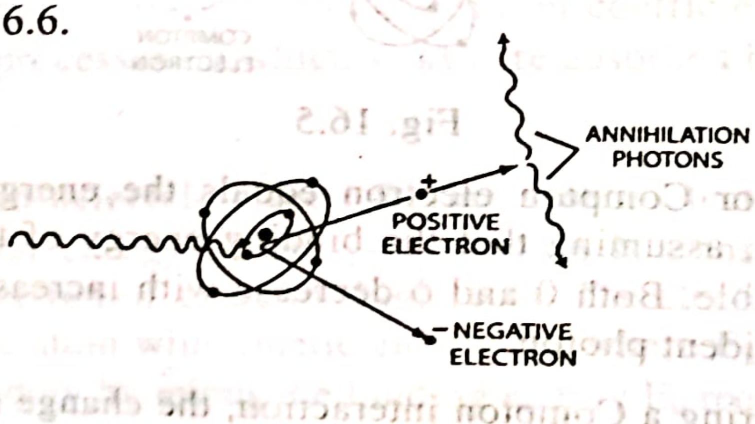


Fig. 16.6

Occasionally, pair production occurs near the coulomb field of an electron rather than that of the nucleus. An interaction near an electron is termed *triplet production,* because the electron receives energy from the photon and is ejected from the atom. Hence, three electrons, two negative and one positive, are released during triplet production. To conserve momentum, the threshold energy for triplet production must be 2.04 MeV. The ratio of triplet to pair production increases with the energy of the incident photons and decreases as the atomic number of the medium is increased.

16.4 Fundamental Laws of Radioactivity

Rutherford and Soddy were the first to give a simple explanation of radioactive disintegration. They argued that when a radioactive disintegration occurs with the emission of α and β particles, the original atom called the *parent* atom changes into something else, called the *daughter.* The daughter atom is also radioactive and disintegrates and changes into a new atom which may also be radioactive. The transformations go on until an inactive substance is reached.

The radioactive transformation mentioned above is governed by a simple law known as the *displacement law* which was discovered by Soddy and Fajans in 1913. The law can be stated as follows:

(1) In all known radioactive transformations either an α or a β particle (*i.e.* never both or more than one of each kind) is emitted by the atom.

(2) When an α particle is emitted by an atom of mass number A and atomic number Z, a new atom is formed whose mass number is less by four units and atomic number less by two units than those of the parent atom *i.e.* the mass number and atomic number of the new atom are (A – 4) and (Z – 2) respectively.

$$_{92}U^{238} \xrightarrow{\alpha} {}_{90}Th^{234}$$

(3) When a β-particle is emitted, the atom has the same mass number A but the atomic number increases by one unit to (Z + 1). This means that in β-decay A does not change but Z and N (number of neutrons) change by one unit in opposite directions.

$$_{90}Th^{234} \xrightarrow{\beta} {}_{91}Pa^{234}$$

The law is readily understood since the alpha particle, identified with the helium nucleus, has mass four times the mass unit and a charge twice the unit charge while the beta particle, identified with the electron, has a very small mass, though a complete unit of negative charge.

Natural Radioactive Series

With the help of the displacement law, one can determine easily the mass and atomic numbers of the different elements in the successive radioactive changes, if the mass and atomic numbers of the parent are known. The first member is called the *parent*, the intermediate members are called *daughters* and the final stable member is called the *end product*. The different atoms produced at the different stages of disintegration will naturally form a series. There are four main series of radioactive elements in nature known at present. (1) Uranium Series (2) Actinium Series (3) Thorium Series and (4) Neptunium Series.

(1) **Uranium Series:** Let us consider the uranium series. The parent element is uranium with mass number A = 238 and Z = 92. It is also known as Uranium I or UI. Uranium I emits α-particles producing Uranium X_1 with Z = 92 – 2 = 90 and A = 238 – 4 = 234. Uranium X_1 emits β-particles and the product formed is Uranium X_2 with A = 234 and Z = 90 + 1 = 91. Uranium X_2 emits β-particles and the product formed is Uranium II with A = 234 and Z = 91 + 1 = 92. U II is an isotope of UI. The sequence of the disintegration is shown below. The end product of the series is stable $_{82}Pb^{206}$

The Uranium Series

Radioactive Species	*Nuclide*	*Type of Disinter-gration*	*Half-Life*	*Disintegration constant* sec^{-1}	*Particle Energy, MeV*
Uranium I	$_{92}U^{238}$	α	4.50 x 10^9 y	4.88 x 10^{-18}	4.20
Uranium X_1	$_{90}TH^{234}$	β	24.1 d	3.33 x 10^{-7}	0.19
Uranium X_2	$_{91}Pa^{234}$	β	1.18 m	9.77 x 10^{-3}	2.32
Uranium Z	$_{91}Pa^{234}$	β	6.7 h	2.88 x 10^{-5}	1.13
Uranium II	$_{92}U^{234}$	α	4.50 x 10^5 y	8.80 x 10^{-14}	4.768
Uranium(Io)	$_{90}TH^{230}$	α	8.0 x 10^4 y	2.75 x 10^{-13}	4.68 m
Uranium(Ra)	$_{88}Ra^{226}$	α	1620 y	1.36 x 10^{-11}	4.777 m
ReEmanation	$_{86}Em^{222}$	α	3.82 d	2.10 x 10^{-6}	5.486
Radium A	$_{84}Po^{218}$	α, β	3.05 m	3.78 x 10^{-3}	α:5.998 β: ?
Radium B	$_{82}Pb^{214}$	β	26.8 m	4.31 x 10^{-4}	0.7
Astatine-218	$_{85}At^{218}$	α	1.5 – 2.0 s	0.4	6.63
Radium C	$_{83}Bi^{214}$	α, β	19.7 m	5.86 x 10^{-4}	α:5.51m β: ?
Radium C'	$_{84}Po^{214}$	α	1.64 x 10^{-4} y	4.23 x 10^3	7.683
Radium C"	$_{81}Tl^{210}$	β	1.32 m	8.75 x 10^{-4}	1.9
RadiumD	$_{82}Pb^{210}$	β	19.4 y	1.13 x 10^{-9}	0.017
RadiumE	$_{83}Bi^{210}$	β	5.0 d	1.60 x 10^{-6}	1.155
RadiumF	$_{84}Po^{210}$	α	138.3 d	5.80 x 10^{-8}	5.300
Thalliu-206	$_{81}Tl^{206}$	β	4.2 m	2.75 x 10^{-3}	1.51
RadiumG	$_{82}Pb^{206}$	Stable			

(2) **Actinium Series:** The parent atom of the series is Actino - uranium with A = 235 and Z = 92. The end product of the series is Actinium D – an isotope of stable lead of atomic mass 207 ($_{82}Pb^{207}$). The series is given below.

The Actinium Series

Radioactive Species	Nuclide	Type of Disintegration	Half-Life	Disintegration constant sec^{-1}	Particle Energy, MeV
Actinouraniu	$_{92}U^{235}$	α	7.10×10^8 y	3.09×10^{-17}	4.559 m
Uranium Y	$_{90}Th^{231}$	β	25.6 h	7.51×10^{-6}	0.30
Protactinium	$_{91}Pa^{231}$	α	3.43×10^4 y	6.40×10^{-13}	5.046 m
Actinium	$_{89}Ac^{227}$	α, β	21.6 y	1.02×10^{-9}	α: 4.94 β:0.046
Radioactinium	$_{90}Th^{227}$	α	18.17 d	4.41×10^{-7}	6.03 m
Actinium K	$_{87}Fr^{223}$	α, β	22 m	5.25×10^{-4}	α: 5.34 β: 1.2
Actinium X	$_{88}Ra^{223}$	α	11.68 d	6.87×10^{-7}	5.864
Astatine-219	$_{85}At^{219}$	α, β	0.9 m	1.26×10^{-2}	α: 6.27
AcEmanation	$_{86}Em^{219}$	α	3.92 s	0.177	6.810 m
Bismuth-215	$_{83}Bi^{215}$	α, β	8 m	1.44×10^{-3}	?
Actinium A	$_{84}Po^{215}$	α, β	1.83×10^{-3} s	3.79×10^{2}	α: 7.37
Actinium B	$_{82}Pb^{211}$	β	36.1 m	3.20×10^{-4}	1.39
Astatine-215	$_{85}At^{215}$	α	10^{-4} s	7×10^{3}	8.00
Actinium C	$_{83}Bi^{211}$	α, β	2.15 m	5.28×10^{-3}	α:6.617m
Actinium C'	$_{84}Po^{211}$	α	0.52 s	1.33	7.442
Actinium C"	$_{81}Tl^{207}$	β	4.79 m	2.41×10^{-3}	1.44
Actinium D	$_{82}Pb^{207}$	Stable			

(3) **Thorium Series: The parent element of this series is Thorium with Z = 90 and A = 232. The series ends with Thorium D – another isotope of stable lead of atomic mass 208 ($_{82}Pb^{208}$). The sequence of disintegration is shown below.**

The Thorium Series

Radioactive Species	Nuclide	Type of Disintergration	Half-Life	Disintegration constant sec^{-1}	Particle Energy, MeV
Thorium	$_{90}Th^{232}$	α	1.39×10^{10}y	1.58×10^{-18}	4.007
Mesothorium 1	$_{88}Ra^{228}$	β	6.7 y	3.28×10^{-9}	0.04
Mesothorium 2	$_{89}Ac^{228}$	β	6.13 h	3.14×10^{-5}	2.18
Radio-thorium	$_{88}Th^{228}$	α	1.910 y	1.15×10^{-8}	5.423
Thorium X	$_{88}Ra^{224}$	α	3.64 d	2.20×10^{-6}	5.681
ThEmanation	$_{86}Em^{220}$	α	51.5 s	1.34×10^{-2}	6.280
Thorium A	$_{84}Po^{216}$	α, β	0.16 s	4.33	6.774
Thorium B	$_{82}Pb^{212}$	β	10.6 h	1.82×10^{-5}	0.58
Astatine-216	$_{85}At^{216}$	α	3×10^{-4} s	2.3×10^{3}	7.79
Thorium C	$_{83}Bi^{212}$	α, β	60.5 m	1.19×10^{-4}	α:6.086m β: 2.25
Thorium C'	$_{84}Po^{212}$	α	3.0×10^{-7} s	2.31×10^{6}	8.780
Thorium C"	$_{81}Tl^{208}$	β	3.10 m	3.73×10^{-3}	1.79
Thorium D	$_{82}Pb^{208}$	Stable			

(4) **Neptunium Series:** The parent atom of this series is Plutonium with Z = 94 and A = 241, but the series is known as Neptunium series because Neptunium has the longest half-life of 2.2×10^{6} years. The mass and the charge number of Neptunium are 237 and 93 respectively.

Before the discovery of transuranic element neptunium, there were only three series as discussed already. The uranium series with uranium as parent element has a mass number = 4n + 2 where n = 58. The actinium series, with the parent element Actinouranium of mass number 235 becomes (4n + 3) series with n = 58. Similarly, Thorium series with the parent element Thorium of mass number 232 in known as 4n series where n = 58.

It was considered that in these series, 4n, (4n+2), and (4n+3), there must be a fourth series (4n+1). This is the Neptunium series. The series ends with Bismuth (A=209). The series is shown below.

Neptunium Series

Substance	*Symbol*	*Atomic Number Z*	*Mass Number A*	*Particles Emitted*	*Half-Life Period T*
Plutonium	Pu	94	241	β	14 years
Americium	Am	95	241	α, γ	470 years
Neptunium	Np	93	237	α	2.2 x 10^6 years
Proto-actinium	Pa	91	233	β	27 days
Uranium	U	92	233	α	1.6 x 10^5 years
Thorium	Th	90	229	β	7340 years
Radium	Ra	88	225	α	15 days
Actinium	Ac	89	225	α	10 days
Francium	Fr	87	221	α	5 minutes
Astatine	At	85	217	α	0.018 second
Bismuth	Bi	83	213	α or β	47 minutes
Thallium	Po	84	213	α	4 x 10^{-6} seconds
Polonium	Tl	81	209	β	2.2 minutes
Lead	Pb	82	209	β	3 hours
Bismuth	Bi	83	209	Stable →	Bismuth A = 209

16.5 Law governing radioactive disintegration

From the very begining it was apparent that radioactivity, whether natural or artificial, is statistical in nature *i.e.*, the rate of disintegration is governed by an *exponential* law. This means that the number of atoms that break up at any instant is not affected by environmental factors (like temperature, pressure, chemical

combination, etc), but is proportional to the number present at that instant. In other words *equal fractions of the radioactive atoms disintegrate in equal interval of time.*

Let N_0 be the number of radioactive atoms present in a sample at the beginning of disintegration, *i.e.*, at t = 0. As time passes, the number of original atoms decreases due to continuous disintegration. At any time t, let the number of radioactive atoms left be N. At this instant the rate of disintegration $(-\frac{dN}{dt})$, also called the *activity*, is proportional to the number present, *i.e.*, N. Or

$$-\frac{dN}{dt} \propto N$$

$$-\frac{dN}{dt} = \lambda N \quad (16.1)$$

where λ is a constant of proportionality known as the *disintegration constant* or *decay constant*. λ is a characteristic of the element that disintegrates, but entirely independent of all external conditions, such as temperature, pressure, etc.

Eqn. (16.1) can be rearranged as

$$\frac{dN}{N} = -\lambda.dt$$

Integrating both sides,

$$\int \frac{dN}{N} = -\lambda \int dt$$

$$\text{or} \quad \log_e N = -\lambda t + C \quad (16.2)$$

C being a constant of integration.

Now at t = 0, N = N_0. Applying this initial condition

$$\log_e N_0 = -\lambda \times 0 + C$$

$$\text{or} \quad C = \log_e N_0$$

Eqn. (16.2) can, therefore, be written as

$$\log_e N = -\lambda t + \log_e N_0$$

or $$\log_e \frac{N}{N_0} = -\lambda t$$

or $$\frac{N}{N_0} = e^{-\lambda t}$$

or $$N = N_0 e^{-\lambda t} \qquad (16.3)$$

Eqn. (16.3) may be used to find the number of the atoms of the original kind still present at any time t.

Radioactive decay constant

Eqn. (16.1) may be written as

$$\lambda = \frac{-dN/dt}{N}$$

Hence, the decay constant may be defined as *the ratio of the number of atoms which disintegrates in unit time to the number of atoms present.*

Let us put $t = \frac{1}{\lambda}$. Then the number of original atoms present after this time is

$$N = N_0 e^{-\lambda \cdot \frac{1}{\lambda}} = N_0 e^{-1} = \frac{N_0}{e} = \frac{N_0}{2.718}$$

or, $N = 0.368\ N_0 \simeq 0.37\ N_0$

Hence the decay constant may also be defined *as the reciprocal of the time during which the number of radioactive atoms falls to 37 percent of its original value.*

16.6 Half-life of a radioactive element

According to eqn. (16.3), the number of radioactive atoms decreases rapidly at first, then more and more slowly, as time goes

on (Fig. 16.7). According to this equation, an infinite time is required, theoretically speaking, for the radioactivity to disappear completely. All radioactive elements are same in this respect. Hence

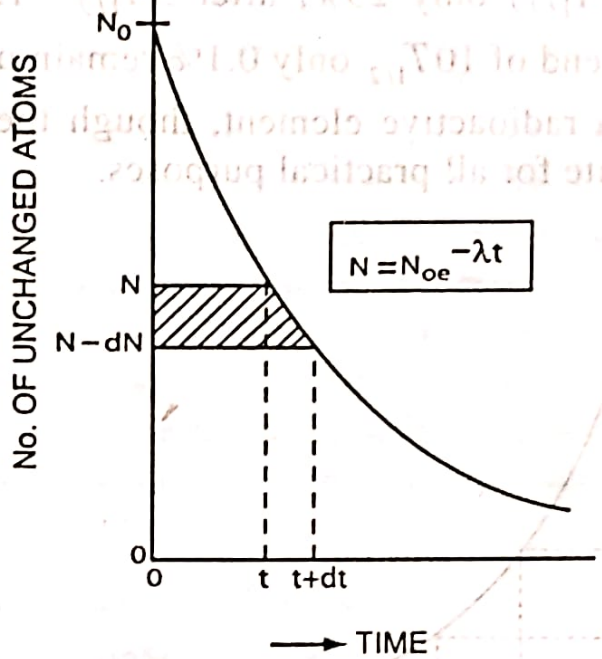


Fig. 16.7

in order to be able to compare one radioactive element with another, a term *half-life* (half-period or simply period) is often used. It is defined *as the time in which the radioactive atoms are reduced to half their initial amount*. It is estimated as follows:

Let the half-life be denoted by $T_{1/2}$. Then according to definition

$$\frac{N}{N_0} = \frac{1}{2} = e^{-\lambda T_{1/2}}; \quad \text{or,} \quad e^{\lambda T_{1/2}} = 2$$

$$\text{or,} \quad \lambda T_{1/2} = \log_e 2; \quad \text{or,} \quad T_{1/2} = \frac{\log_e 2}{\lambda}$$

$T_{1/2}$ is inversely proportional to λ; the greater the value of the decay constant of a radioactive element, the shorter will be its half-life and *vice versa*.

The half-life of radium is 1620 years. This means that it takes 1620 years for one-half of a given quantity of radium to change into its daughter product radon. In another 1620 years, ½ of the

remainder would have disintegrated leaving ¼ of the original amount behind. However, radon has a half-life of only 3.8 days.

Thus for a given radioelement, at the end of one half-life ($T_{1/2}$) 50% of the radioactive atoms remain unchanged, at the end of two half-lives ($2T_{1/2}$) only 25%, after $3T_{1/2}$ 12.5%, after $4T_{1/2}$ 6.25% and at the end of $10T_{1/2}$ only 0.1% remain unchanged. Hence, the life time of a radioactive element, though theoretically infinite (Fig. 16.8), is finite for all practical purposes.

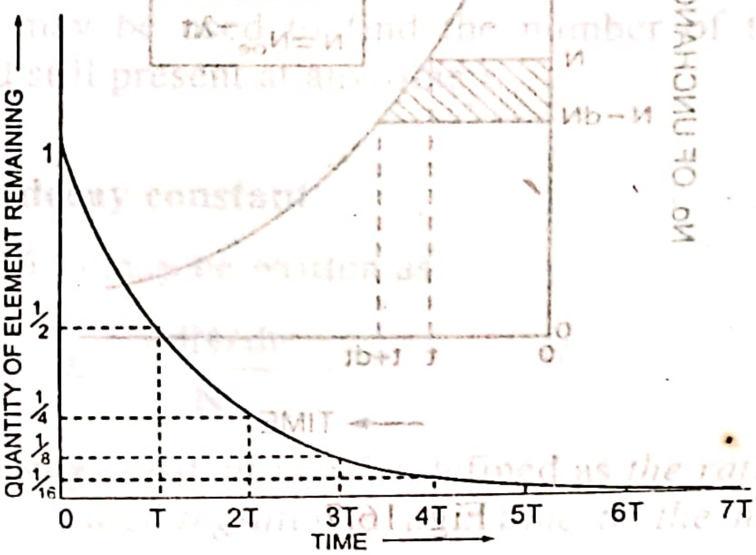


Fig. 16.8

As can be seen from eqn. (16.1), λN represents the activity *i.e.*, the rate at which N atoms disintegrate.

$$\therefore \quad \text{rate of disintegration or activity} = \lambda N = \frac{0.693}{T_{1/2}} N$$

If $T_{1/2}$ is expressed in seconds, then the above equation gives the emission rate in terms of the number of particles expelled per second from N atoms of the given nuclide.

16.7 Mean life (or average life) of a radioactive element

The statistical law governing radioactive decay states nothing about the decay of individual radioactive atoms. Among the disintegrating atoms, some have only a short existence, while others

remain unchanged for a long time – why we do not know. At a given moment why a particular atom disintegrates and not its neighbour, also remains inexplicable. The past history of an atom, chiefly the fact that it may have remained stable for a long time, seems to have no influence whatsoever on the moment of its disintegration.

Since the radioactive atoms are constantly disintegrating one after another, the life of every atom is different. The atoms which disintegrate earlier have very short life whereas others which disintegrate at the end have a very long life. Thus the possible time of existence of a radioactive atom may vary from zero to infinity and all radioactive atoms are same in this respect. As in the case of all statistical problems, here also a quantity known as the *mean* or *average life* (τ) may be employed to differentiate radioelements. The mean life of a radioactive element is defined *as the ratio of the total life time of all the radioactive atoms to the total number of such atoms in it.*

$$\text{or} \qquad \tau = \frac{\text{sum of lives of all atoms}}{\text{total number of atoms}}$$

Let N_0 be the total number of radioactive atoms in the beginning. Let N be the number of atoms of that element after time t. Then

$$N = N_0 e^{-\lambda t}$$

$$\text{or,} \qquad -\frac{dN}{dt} = \lambda N$$

$$\therefore \qquad dN = \lambda\, N\, dt$$

The negative sign has been left out. Negative sign merely indicates that dN atoms disintegrate in the very short interval dt. The number of such atoms itself is a positive quantity.

Now the dN atoms in question have had a life time between t and (t + dt). Since dt is very small, we may assume without serious error that all these atoms have lived for a time t. Hence the total life time of these dN atoms is equal to t.dN.

The possible life of any of the total number of atoms varies from 0 to ∞, the total life-tiem N_0 atoms is given by

$$\text{Now, the mean life } \tau = \frac{\int_{t=0}^{t=\infty} t.dN}{\text{total number of atoms}} \quad \frac{\text{total life time}}{\text{total number of atoms}}$$

$$= \frac{\int_0^\infty t.dN}{N_0}$$

Substituting for dN,

$$\tau = \frac{\int_0^\infty t.\lambda.N.dt}{N_0} = \frac{\int_0^\infty t.\lambda.N_0 e^{-\lambda t}.dt}{N_0}$$

$$= \lambda \int_0^\infty t.e^{-\lambda t}.dt$$

Integrating by parts, let $t = u$, so that $dt = du$ and $e^{-\lambda t}.dt = dv$, which gives $v = e^{-\lambda t}/-\lambda$; then

$$\tau = \lambda \left[vu - \int vdu\right]_0^\infty$$

$$= \lambda \left[t.\frac{e^{-\lambda t}}{-\lambda} - \int \frac{e^{-\lambda t}}{-\lambda} dt\right]_0^\infty$$

$$= \lambda \left[t\frac{e^{-\lambda t}}{-\lambda} - \frac{e^{-\lambda t}}{\lambda^2}\right]_0^\infty$$

$$= \lambda \left[\frac{(-\lambda t e^{-\lambda t}) - (-e^{-\lambda t})}{\lambda^2}\right]_0^\infty$$

$$= -\frac{1}{\lambda}\left[(\lambda t + 1)e^{-\lambda t}\right]_0^\infty$$

$$= \frac{1}{\lambda}$$

Thus the mean life τ turns out to be the reciprocal of λ.

From $\quad T_{1/2} = \dfrac{0.693}{\lambda}$, we get

$$\tau = \frac{T_{1/2}}{0.693} \cong 1.44 T_{1/2}$$

or $\quad T_{1/2} = 0.693\tau$.

The mean life of radium is therefore equal to 1620/0.693 = 2337.66 years and that of radon 3.8/0.693 = 5.5 days.

***Example 16.1** The half-life of a radioactive substance is 30 days. Calculate (i) the radioactive decay constant, (ii) the mean life (iii) the time taken for ¾ of the original number of atoms to disintegrate and (iv) the time for 1/8 of the original number of atoms to remain unchanged.*

Soln:

(i) $\quad T_{1/2} = 30$ days

$$\therefore \quad \lambda = \frac{0.693}{T_{1/2}} = \frac{0.693}{30\,d} = 0.0231 \text{ per day}$$

(ii) $\quad$ Mean life, $\tau = \dfrac{1}{\lambda} = \dfrac{1}{0.0231 d^{-1}} = 43.29$ days

(iii) From $N = N_0 e^{-\lambda t}$, we have

$$\frac{1}{4}N_0 = N_0 e^{-\lambda t} \quad \text{where} \quad N = N_0 - \frac{3N_0}{4} = \frac{1}{4}N_0$$

or $\quad \dfrac{\frac{1}{4}N_0}{N_0} = e^{-\lambda t}$; or $e^{-\lambda t} = \dfrac{1}{4}$

or $\quad e^{\lambda t} = 4 \quad ; \quad \lambda t = \log_e 4$

$$\therefore \quad t = \frac{\log_e 4}{\lambda} = \frac{}{0.0231} = 60 \text{ days}$$

(iv) The number of atoms left, $N = \frac{1}{8} N_0$

$$\therefore \quad \frac{N}{N_0} = \frac{1}{8} = e^{-\lambda t}$$

or $\quad e^{\lambda t} = 8$

$$\therefore \quad \lambda t = \log_e 8 \;\; ; \text{ or } \; t = \frac{\log_e 8}{0.0231} = \frac{}{0.0231} = 90 \text{ days}$$

Example 16.2 The half-life of radium is 1620 years. In how many years will one gram of pure element (i) lose one centigram and (ii) be reduced to one centigram?

Soln.

The decay constant of radium is

$$\lambda = \frac{0.693}{T_{1/2}} = \frac{0.693}{1620 \text{ years}} = 4.28 \times 10^{-4} \text{ y}^{-1}$$

(i) Let t be the time during which one centigram of radium is lost due to disintegration. The amount remaining is $(1 - \frac{1}{100}) = 0.99$ gm

From $\quad N = N_0 e^{-\lambda t}$, we have

$$\frac{N}{N_0} = e^{-\lambda t}; \text{ or } \frac{0.99 N_0}{N_0} = e^{-\lambda t}$$

or $\quad e^{-\lambda t} = 0.99 = \frac{99}{100}$

$$e^{\lambda t} = \frac{100}{99}; \text{ or } \lambda t = \log_e(\frac{100}{99});$$

$$\therefore \quad t = \frac{\log_e(100/99)}{\lambda} = \frac{}{4.28\times10^{-4}\ \text{years}} = 23.68 \text{ years}$$

(ii) Now N = 0.01 gm

$$\therefore \quad \frac{N}{N_0} = 0.01 = \frac{1}{100} = e^{-\lambda t}$$

or $\quad e^{\lambda t} = 100 \qquad \therefore \lambda t = \log_e 100$

$$\therefore \quad t = \frac{\log_e 100}{\lambda} = \frac{}{4.28\times10^{-4}\ \text{years}} = 10{,}760 \text{ years}$$

***Example 16.3** 1gram of radium is reduced by 2.1 mg in 5 years by α-decay. Calculate the half-life of radium.*

Soln.

The amount of radium left at the end of 5 years is, $N = 1 - 2.1\times10^{-3}$ $= 1 - 0.0021 = 0.9979$ gm

From $\frac{N}{N_0} = e^{-\lambda t}$ we have

$$\frac{0.9979}{1.0} = e^{-\lambda t} = 0.9979$$

Now t = 5 years

$$e^{-5\lambda} = 0.9979 \ ; \quad \text{or} \quad e^{5\lambda} = \frac{1}{0.9979}$$

or $\quad 5\lambda = \log_e\left(\frac{1}{0.9979}\right)$

$$\therefore \quad \lambda = \frac{\log_e(1/0.9979)}{5\text{ y}} = \frac{}{5\text{ y}} = 41.4468\times10^{-5}\text{ y}$$

$$\therefore \quad T_{1/2} = \frac{0.6931}{41.4468\times10^{-5}\text{ y}} = 1672 \text{ years.}$$

Example 16.4 A radioactive source contains one microgram (μg) of Pu – 239. This source is estimated to emit 2200 α-particles per second in all directions. Estimate the half-life of plutonium from this data.

Soln.

Mass of one α-particle = 4.00387 amu

$= 4.00387 \times 1.66 \times 10^{-24}$ gm $= 6.646 \times 10^{-24}$ g

$\therefore$ Mass of 2200 α-particles $= 2200 \times 6.646 \times 10^{-24}$ gm

$= 14621.2 \times 10^{-24}$ gm

$$\text{Now } \lambda = \frac{-dM/dt}{M} = \frac{14621.2 \times 10^{-24}/1}{1 \times 10^{-6}} = 14621.2 \times 10^{-18} s^{-1}$$

$$\therefore \quad \text{Half-life of Pu, } T_{1/2} = \frac{0.693}{\lambda} = \frac{0.693}{14621.2 \times 10^{-18}} \text{ seconds}$$

$= 4.74 \times 10^{13}$ seconds $= 15 \times 10^{5}$ years.

Example 16.5 A radioactive sample initially contains 4.00 mg of ${}_{92}U^{234}$ (a) How much of it will remain unchanged after 62,000 years? (b) What will be the activity of ${}_{92}U^{234}$ at the end of that time? $T_{1/2} = 2.48 \times 10^{5}$ years and $\lambda = 8.88 \times 10^{-14}\ s^{-1}$.

Soln.

(a) Let M_0 and M be the masses of the radioactive material at the begining and the end of 62,000 years.

Now $M = M_0 e^{-\lambda t}$ $\qquad$ Here, $\lambda t = \dfrac{0.693}{T_{1/2}} \times t$

$\therefore \quad M = M_0 e^{-0.173}$ $\qquad$ $= \dfrac{0.693}{2.48 \times 10^{5}\ y} \times 62{,}000\ y$

$\dfrac{M}{M_0} = e^{-0.173}$ $\qquad$ $= 0.173$

or $$e^{0.173} = \frac{M_0}{M} = 1.189$$

$$[\lambda t = 8.88\times10^{-14}\times62000\times365\times24\times3600 = 0.173]$$

or $$M_0 = 1.189\ M$$

or $$M = \frac{M_0}{1.189} = \frac{4}{1.189}\text{mg} = 3.364\text{ mg}$$

So mass left unchanged = 3.364 mg

(b) Activity is given by

$$A = \frac{dN}{dt} = \lambda N$$

where N is the number of atoms in 3.364 mg

$$\text{Now } N = \frac{3.364\times10^{-3}\times6.02\times10^{23}}{234}$$

where 234 is the atomic weight of the radioactive substance and 6.02×10^{23} is the Avogadro's number.

$$\text{Activity} = \lambda N = 8.8\times10^{-14}\times\frac{3.364\times10^{-3}\times6.02\times10^{23}}{234}$$

$$= 7.69\times10^{5}\text{ disintegrations per seconds.}$$

***Example 16.6** A counter rate meter is used to measure the activity of a radioactive sample. At a certain instant, the count rate was recorded as4750 counts per minute. Five minutes later, the count rate recorded was 2700 counts per minute. Compute (i) the decay constant and (ii) the half-life of the sample.*

Soln.

(a) $$N = N_0 e^{-\lambda t}$$

or $$\frac{N}{N_0} = e^{-\lambda t}\ ;\ e^{\lambda t} = \frac{N_0}{N}$$

Now $N_0 = dN_1/dt = 4750$

and $N = dN_2/dt = 2700$

$$\therefore \quad e^{5\lambda} = \frac{4750}{2700}$$

$$\text{or} \quad 5\lambda = \log e \left(\frac{4750}{2700}\right)$$

$$\text{or} \quad \lambda = \frac{\log_e(1.76)}{5} = 0.113 \text{ per minute.}$$

(b) $$T_{1/2} = \frac{0.693}{0.113} = 6.1 \text{ minutes.}$$

16.8 Units of Radioactivity

In radioactivity the number of radioactive atoms which disintegrate in unit time *i.e.*, dN/dt is called the *activity* of the substance. It is this activity rather than the total amount of the substance expressed by weight or in number of atoms which is of real importance. The unit for measuring activity is called the *Curie (Ci)* which is defined *as that quantity of a radioactive substance which gives* 3.7×10^{10} *disintegrations per second.*

Obviously, 1 millicurie (mCi) = 3.7×10^{10} dis/sec

1 microcurie (μCi) = 3.7×10^{4} dis/sec.

The SI unit of activity is *Becquerel (Bq)* which is equal to 1 *disintegration per second*

1 MBq = 10^6 Bq and 1 GBq = 10^9 Bq.

Another unit, called *Rutherford (Rd)*, is also used to measure activity. It is defined as the quantity of a radioactive substance which gives 10^6 disintegrations per second.

Smaller units like millirutherford (mrd) and microrutherford (μrd) are also used.

It is to be noted that for a substance with very short half-life, very little of the substance is required for 1 Curie of activity. On the

other hand, for the substance with very long half-life, a very large quantity of the substance is required for 1 curie of activity.

Let us find out the mass in gm of radium-B *i.e.*, RaB ($_{82}Pb^{214}$) that will give 1 Ci and 1 rd of activity. The half-life of RaB is 26.8 minutes. So its decay constant

$$\therefore \quad \lambda = \frac{0.693}{26.8\times 60} = 4.31\times 10^{-4}\ s^{-1}$$

Let m gm be the mass of RaB that gives 1 Ci of activity *i.e.*, 3.7×10^{10} disintegrations per second.

The number of atoms contained in m gm

$$= \frac{m\times 6.02\times 10^{23}}{214}$$

Now activity, $-\frac{dN}{dt} = \lambda N$

$$\therefore \quad 3.7\times 10^{10} = \frac{4.31\times 10^{-4}\times m\times 6.02\times 10^{23}}{214}$$

$$\text{or,} \quad m = \frac{214\times 3.7\times 10^{-10}}{4.31\times 10^{-4}\times 6.02\times 10^{23}} = 3.1\times 10^{-8}\ \text{gm.}$$

Similarly for 1 rd, $\frac{dN}{dt} = 10^6$ disintegrations per second

$$\therefore \quad 10^6 = 4.31\times 10^{-4}\times \frac{m\times 6.02\times 10^{23}}{214}$$

$$\text{or} \quad m = 8.3\times 10^{-13}\ \text{gm}$$

It is, therefore, obvious that for a substance of short half-life, a very small amount of material is required to provide one curie or one rutherford of activity.

Let us now consider what amount of $_{92}U^{238}$, which has a very long half-life of 4.5×10^9 years, will give 1 Ci or 1 rd of activity.

The decay constant of U^{238}.

$$\lambda = \frac{0.693}{4.5\times10^{9}\times365\times24\times3600} = 4.9\times10^{-8}\ s^{-1}$$

Using $-\frac{dN}{dt} = N\lambda$, we get

for one curie

$$3.7 \times 10^{10} = \frac{4.9\times10^{-18}\times m\times6.02\times10^{23}}{238}$$

$\therefore \quad m = 3.2 \times 10^{6}$ gm $= 3.2 \times 10^{3}$ kg

$= 3.2$ metric ton.

for one Rutherford

$$10^{6} = \frac{4.9\times10^{-18}\times m\times6.02\times10^{23}}{238}$$

$\therefore \quad m = 80.9$ gm

Example 16.7 Find the activity of 1mg (10^{-3}gm) of radon (R_n^{222}). The half-life of radon is 3.8 days.

Soln.

$$\lambda = \frac{0.693}{3.8\times24\times3600} = 2.1\times10^{-6}\ s^{-1}$$

Number of atoms in 10^{-3} gm,

$$N = \frac{10^{-3}\times6.02\times10^{23}}{222}$$

So activity, $A = \frac{dN}{dt}$ (ignoring the minus sign)

$$= \lambda N$$

$$= \frac{2.1\times10^{-6}\times10^{-3}\times6.02\times10^{23}}{222}$$

$$= 5.7 \times 10^{12} \text{ disintegration per second}$$

$$= \frac{5.7\times10^{12}}{3.7\times10^{10}} = 153 \text{ Ci}$$

$$= \frac{5.7\times10^{12}}{10^{6}} = 5.7 \times 10^{6} \text{ rd}$$

$$= 5.7 \times 10^{3} \text{ GBq.}$$

16.9 Isotopes, Isobars, Isotones, Isodiapheres and Isomers

Isotopes: Atoms which have the same atomic number Z (hence have similar chemical properties) but different mass number A (atomic weight) are called *isotopes* (meaning the same place in the periodic table). Obviously, the different isotopes of an element contain different number of neutrons in their nuclei.

Example:

$$_{29}Cu^{63} \text{ and } _{29}Cu^{65} \; : \; _{92}U^{238} \text{ and } _{92}U^{235}$$

$$_{90}Th^{232} \text{ and } _{90}Th^{228} \; : \; _{82}Pb^{214} \text{ and } _{82}Pb^{210}$$

Isobars: Atoms which have the same mass number A but different atomic number Z are called *isobars*. Obviously, these atoms belong to different chemical elements.

Example:

$$_{18}Ar^{40} \text{ and } _{20}Ca^{40} \; ; \; _{38}Sr^{92} \text{ and } _{42}Mo^{92}$$

$$_{24}Cr^{54} \text{ and } _{26}Fe^{54} \; ; \; _{32}Ge^{76} \text{ and } _{34}Se^{76}$$

Isotones: Atoms which have the same neutron numbers [N = A - Z] are called *isotones*.

Example:

$_{17}Cl^{37}$ and $_{19}K^{39}$ are isotones since each has 20 neutrons in its nucleus.

Isomers: These are atoms (nucleus), which have the same mass number A and the same atomic number Z, but disintegrates in different ways with different time periods. This manifestation of different radioactive properties argues to certain differences in internal structure of nuclei which are otherwise identical in all respects. They are called nuclear isomers.

Example:

${}_{91}UZ^{234}$ and ${}_{91}UX_2^{234}$ are isomeric. UX_2 has a half-life of 1.18 minutes and it emits three groups of β-rays with different end-point energies. On the other hand UZ has a half-life of 6.7 hours and emits four group of β-rays with different end-point energies.

Similarly ${}_{35}Br^{80}$ is also isomeric – one of the member having a half-life of 6.4 minutes while the other has a half-life of 18 minutes.

Isodiapheres: Atoms having the same excess number of neutrons over protons are called *isodiapheres.*

Example:

${}_{17}Cl^{37}$ and ${}_{18}Ar^{39}$ are isodiapheres as the number of neutrons in each exceeds the number of protons by three.

In general, a nuclide of atomic number Z and mass number A has A nucleons of which Z are protons and (A – Z) are neutrons. Nuclides having the same value of Z are isotopes, those having the same value of A are isobars, those having the same value of (A – Z) are isotones; those having the same value of (A–2Z) are isodiapheres and those having the same values of Z and A but different radioactive properties are isomers.

For example, ${}_{17}Cl^{37}$ is an isotope of ${}_{17}Cl^{35}$, an isobar of ${}_{16}S^{37}$, an isotone of ${}_{19}K^{39}$ and an isodiapheres of ${}_{18}Ar^{39}$.

16.10 Law of successive disintegration and radioactive equilibrium

In radioactive disintegration, the original atoms disappear giving rise to new atoms (atoms of new element). These new atoms are also radioactive leading to a long chain of different

radioactive atoms in the form of a series. The transformations go on until an inactive *i.e.*, stable substance is reached. For example, consider a substance A which decays to form a substance B. B then decays to form a substance C and so on.

$$A \longrightarrow B \longrightarrow C \longrightarrow \ldots\ldots X \text{ (stable)}$$

In a radioactive series, any two adjacent elements may be considered as *parent* and *daughter*, the former being that which by its own decay produces the later. As for example, radium, which emits an α-particle and produces radon, is the parent with respect to radon, the daughter. Obviously in a *series* disintegration, the parent of the following element will be the daughter of the preceding element. The law governing successive disintegrations deals with the quantity of the daughter substance present at any time and determines the condition of equilibrium between the daughter and the parent.

We shall here consider the simplest case of two radioactive elements. Cases involving more than two elements can be similarly treated but in a more general way.

The daughter, as it is being produced from the parent according to an exponential law, will, at the same time, also decay exponentially. Naturally, the formation of the daughter will depend on the decay constant of the parent, while its own decay will depend on its own decay constant.

Let N_1 be the amount of the parent substance at any instant t and λ_1 its decay constant, while N_2 be the amount of the daughter at the same instant and λ_2 its decay constant. Let N_1^0 and N_2^0 be the initial amount of parent and daughter substance respectively at $t = 0$.

The rate at which the parent decays

$$\frac{dN_1}{dt} = -\lambda_1 N_1$$

Since an atom of the daughter appears every time an atom of the parent disappears, the rate of formation of the daughter is $\lambda_1 N_1$. At the same time the daughter is also disappearing according to

its own decay constant and the rate of disappearance of the daughter is obviously $\lambda_2 N_2$. So the rate of net increase of the daughter at the same time is given by

$$\frac{dN_2}{dt} = \lambda_1 N_1 - \lambda_2 N_2$$

Hence $$\frac{dN_2}{dt} = \lambda_1 N_1^0 e^{-\lambda_1 t} - \lambda_2 N_2 \quad [N_1 = N_1 e^{-\lambda_1 t}]$$

$$\therefore \quad \frac{dN_2}{dt} + \lambda_2 N_2 = \lambda_1 N_1^0 e^{-\lambda_1 t}$$

Multiplying both sides by the integrating factor $e^{\lambda_2 t}$,

$$\frac{dN_2}{dt} e^{\lambda_2 t} + \lambda_2 N_2 e^{\lambda_2 t} = \lambda_1 N_1^0 e^{-\lambda_1 t} e^{\lambda_2 t}$$

$$= \lambda_1 N_1^0 e^{(\lambda_2 - \lambda_1)t}$$

or, $$\frac{d}{dt}[N_2 e^{\lambda_2 t}] = \lambda_1 N_1^0 e^{(\lambda_2 - \lambda_1)t}$$

Integrating

$$N_2 e^{\lambda_2 t} = \frac{\lambda_1}{\lambda_2 - \lambda_1} \lambda_1 N_1^0 e^{(\lambda_2 - \lambda_1)t} + C \quad (16.4)$$

where C is a constant of integration and can be determined from the initial condition.

Let us assume that $N_2^0 = 0$ at time $t = 0$, *i.e.*, the initial amount of the daughter is zero.

Putting this in equation (16.4), $N_2 = N_2^0 = 0$ at $t = 0$

or $$C = -\frac{\lambda_1}{\lambda_2 - \lambda_1} N_1^0$$

Substituting this value of C in eqn. (16.4)

$$N_2 e^{\lambda_2 t} = \frac{\lambda_1 N_1^0}{\lambda_2 - \lambda_1}[e^{(\lambda_2-\lambda_1)t} - 1]$$

or, $$N_2 = \frac{\lambda_1 N_1^0}{\lambda_2 - \lambda_1}[e^{-\lambda_1 t} - e^{-\lambda_2 t}] \qquad (16.5)$$

If we cannot assume that the initial amount of daughter is not zero, *i.e.*, N_2^0 is not zero at t = 0, then the above equation becomes somewhat more complicated and is given by

$$N_2 = N_1^0 \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} + (N_2^0 - \frac{N_1^0 \lambda_1}{\lambda_2 - \lambda_1}) e^{-\lambda_2 t} \qquad (16.6)$$

However, for our study we shall use the simplified eqn. (16.5) which assumes that initially only the parent atom alone is present.

We shall now consider some important cases:

Case 1: Transient equilibrium

When the half-life of the parent nuclide is not much longer than the half-life of the daughter, *i.e.*, the mean lives of the parent and daughter are of the same order of magnitude ($\lambda_1 \approx \lambda_2$), it can be shown that in such a case the daughter substance first reaches a maximum and then decreases at the decay rate of the longer lived of the two, *i.e.*, either of the parent or the daughter according to whether $\lambda_1$is smaller or longer than λ_2.

The time in which the daughter reaches a maximum can be obtained by differentiating eqn. (16.5) and setting it to zero. Thus

$$\frac{dN_2}{dt} = A(e^{-\lambda_1 t_{max}} + \lambda_2 e^{-\lambda_2 t_{max}}) = 0$$

where $A = N_1^0 \frac{\lambda_1}{\lambda_2 - \lambda_1}$

Since A cannot be equal to zero,

$$-\lambda_1 e^{-\lambda_1 t_{max}} + \lambda_2 e^{-\lambda_2 t_{max}} = 0$$

or $$\lambda_2 e^{-\lambda_2 t_{max}} = \lambda_1 e^{-\lambda_1 t_{max}}$$

i.e., $$\frac{\lambda_2}{\lambda_1} = \frac{e^{\lambda_2 t_{max}}}{e^{\lambda_1 t_{max}}}$$

or, $$\log_e\left(\frac{\lambda_2}{\lambda_1}\right) = (\lambda_2 - \lambda_1)t_{max}$$

$$\therefore \quad t_{max} = \frac{1}{(\lambda_2 - \lambda_1)}\log_e \frac{\lambda_2}{\lambda_1} \qquad (16.7)$$

After this time (t_{max}), the daughter, will have a decay rate dependent on the relative values of λ_1 and λ_2.

If $\lambda_1 < \lambda_2$, *i.e.*, the parent has a longer half-life than the daughter, the second term within the bracket of eqn.(16.5), *viz* $e^{-\lambda_2 t}$ will approach zero faster than the first, so that N_2 will eventually break down at the decay rate of the parent. Eqn. (16.5) then reduces to

$$N_2 = N_1^0 \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}$$

$$= N_1 \frac{\lambda_1}{\lambda_2 - \lambda_1}$$

$$\therefore \quad \frac{N_2}{N_1} = \frac{\lambda_1}{\lambda_2 - \lambda_1} \qquad (16.8)$$

Hence both the substances will eventually decay with the half-life of the parent and their ratio N_2/N_1 remains constant. This state of affairs is called *transient equilibrium*.

For example Te^{132} ($T_{1/2} = 78$ hr) decays into I^{132}($T_{1/2}$ = 2.3 hr). The transient equilibrium for this transition is illustrated in Fig. 16.9. The

activity of I^{132} becomes greatest when parent (Te^{132}) and daughter I^{132} activities are equal. At all later times, the daughter activity exceeds the

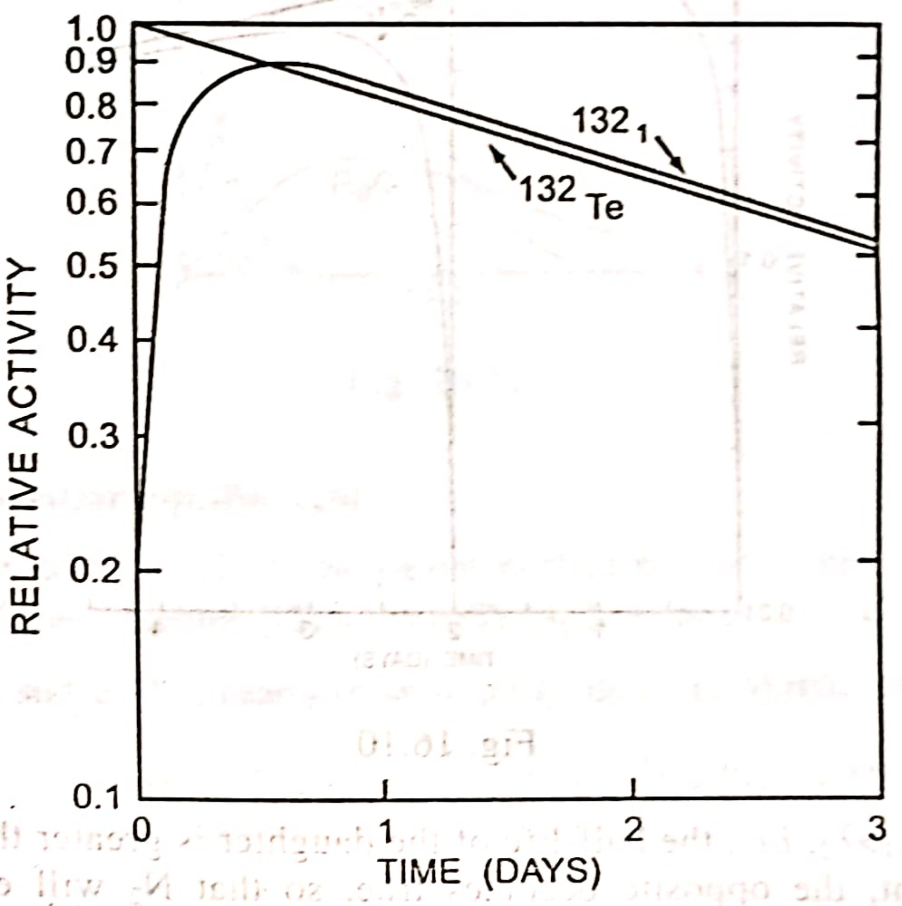


Fig. 16.9

activity of the parent, and both nuclides decay with the half-life of the parent. After the moment of transient equilibrium, the ratio N_2/N_1, remains constant as given by eqn. (16.8).

The principle of transient equilibrium underlines the production of short-lived isotopes (*e.g.*, Te^{99m} and Is^{113m}) in generator for use as radio-active nuclides in nuclear medicine. For example, the activities of Mo^{99} ($T_{1/2}$ = 66 hr) and (Tc^{99m} $T_{1/2}$ = 6 hr) are plotted in Fig. 16.10 as a function of time. The Tc^{99m} activity remains less than that for Mo^{99} because about 14% of the Mo^{99} decay promptly to Tc^{99} without passing through the isomeric state of Tc^{99m}. The abrupt decrease in activity at 48 hr reflects removal of Tc^{99m} from the generator.

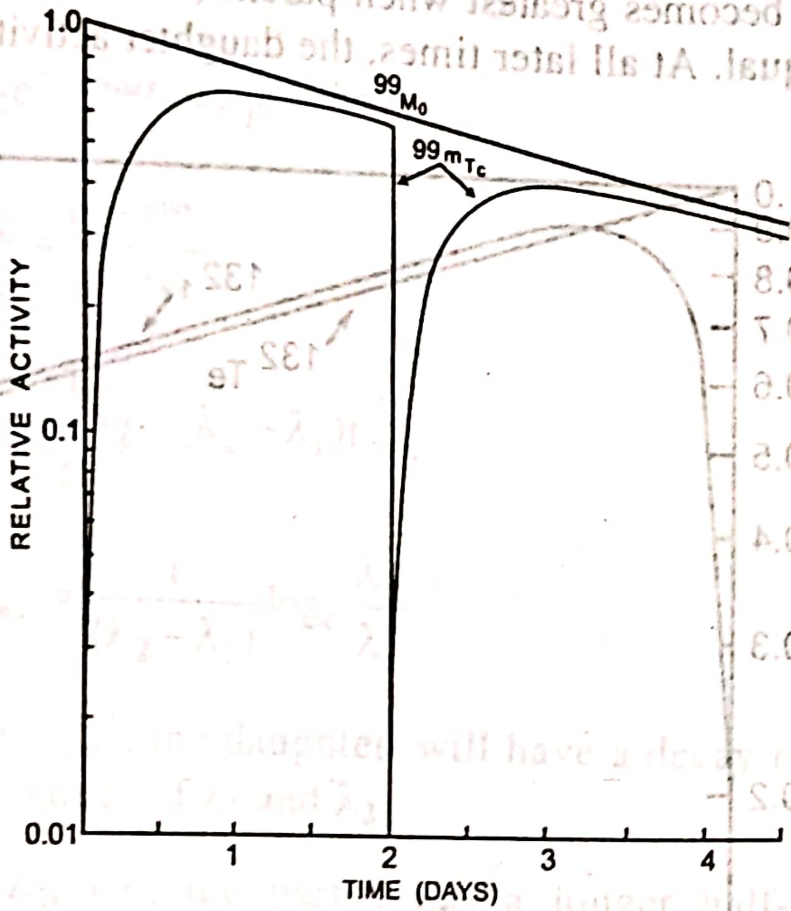


Fig. 16.10

If $\lambda_1 > \lambda_2$, *i.e.*, the half-life of the daughter is greater that that of the parent, the opposite becomes true, so that N_2 will eventually decay with its own decay constant as eqn. (16.5) now reduces to

$$N_2 = N_1^0 \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}$$

After a certain time the parent substance practically disappears and there remains only the daughter which breaks at its own rate. The decay of the active deposits of radium provides an illustration of this. The decay constant of the active deposits of radium are

R_aA ($\lambda = 3.97 \times 10^{-3}$)

R_aB ($\lambda = 4.31 \times 10^{-4}$)

R_aC ($\lambda = 5.86 \times 10^{-4}$)

Since the decay constant of R_aA is greater than those of R_aB and R_aC, these two daughters will eventually decay at their own rates, the parent R_aA having practically disappeared as seen from the decay curves of Fig. 16.11.

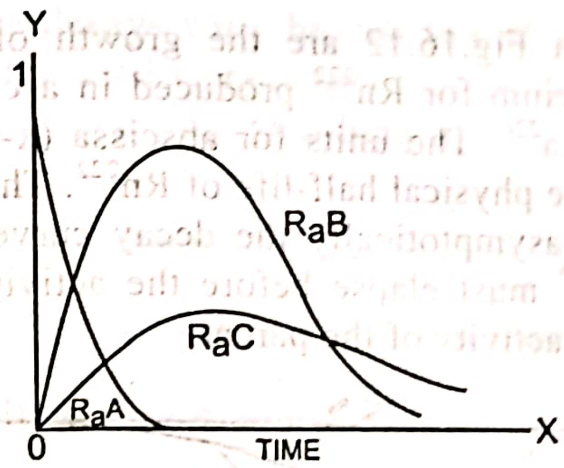


Fig. 16.11

Case II: Secular equilibrium

When the half-life of the parent nuclide is much greater than the half-life of the daughter $[(T_{1/2})_1 >> (T_{1/2})_2]$, $\lambda_1 << \lambda_2$, $\lambda_2 - \lambda_1$ is nearly equal to λ_2 and $e^{-\lambda_1 t}$ is nearly equal to unity. Eqn. (16.5) reduces to

$$N_2 = N_1^0 \frac{\lambda_1}{\lambda_2 - \lambda_1}(e^{-\lambda_1 t} - e^{-\lambda_2 t}) \cong \frac{\lambda_1}{\lambda_2} N_1^0 (1 - e^{-\lambda_2 t})$$

After several half-lives of the daughter have elapsed, $e^{-\lambda_2 t} = 0$. Then

$$N_2 = N_1^0 \left(\frac{\lambda_1}{\lambda_2}\right) \ldots \ldots \ldots \text{(i)}$$

$$\lambda_2 N_2 = \lambda_1 N_1^0$$

From eqn. (i), it is obvious that the amount of daughter present is practically constant, having the value $N_2 = N_1^0 (\frac{\lambda_1}{\lambda_2})$. It is then that the daughter is said to be in *secular* or *permanent equilibrium* with the parent. Since the amount of parent is constant, $N_1^0 = N_1$ and $N_2/N_1 = \lambda_1/\lambda_2 = \tau_2/\tau_1$. The amounts of the two substances present are inversely proportional to their decay constant or directly proportional to their respective mean lives. Under these conditions the daughter breaks up as fast as it is formed.

Illustrated in Fig.16.12 are the growth of activity and the activity at equilibrium for Rn^{222} produced in a closed environment by the decay of Ra^{226}. The units for abscissa (x-axis) in the figure are multiples of the physical half-life of Rn^{222}. The growth curve for Rn^{222} approaches asymptotically the decay curve for Rn^{226}. Seven half-lives of Rn^{222} must elapse before the activity of the daughter equals 99% of the activity of the parent.

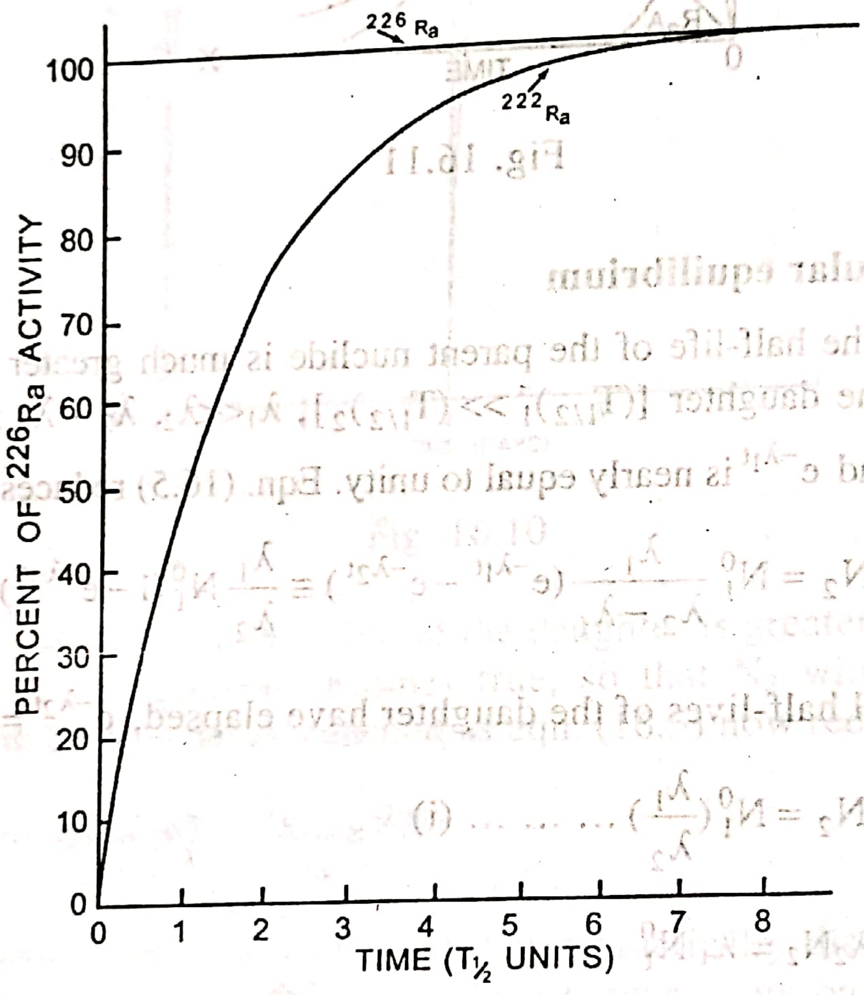


Fig. 16.12

If the half-life of the parent is less than that of the daughter [$(T_{1/2})_1 < (T_{1/2})_2$ or $\lambda_1 > \lambda_2$], then a constant relationship is not achieved between the activities of the parent and the daughter. Instead, the activity of the daughter increases initially, reaches a maximum, and then decreases with a half-life intermediate between the half-lives of the parent and the daughter. Radioactive nuclides in secular equilibrium has important medical application. For example, energetic beta particles from Y^{90} in secular equilibrium with Sr^{90} are used to treat intraocular lesions. The Sr^{90} - Y^{90} opthalmic irradiator decays with the physical half-life of Sr^{90} (28 years), whereas a

source of Y^{90} alone decays with the half-life of Y^{90} (64 hr). Radium needles and capsules used in radiation therapy contain many decay products in secular equilibrium with long-lived Rn^{226}.

16.11 Measurements of decay constants

The three constants λ, $T_{1/2}$ and τ of a radioelement are inter-related by the relations $T_{1/2} = \frac{0.693}{\lambda}$ and $\tau = \frac{1}{\lambda}$. If one of them is experimentally measured, the other two can be readily deduced. Usually the decay constant λ is determined experimentally. Different methods are employed for the determination of λ depending on whether λ is large or small.

(i) Elements of short half-life (large λ).

The decay curve is drawn from the measured values of intensity with time. If the logarithm of intensity is plotted against time, a straight line graph is obtained representing the relation

$$\log N = \log N_0 - \lambda t$$

or $$y = a - \lambda t$$

where $\log N = y$ and long $N_0 = a$ (constant).

$$\therefore \quad \frac{dy}{dt} = -\lambda$$

Hence the slope of the straight line graph gives the value of λ.

(ii) Elements of long half-life (small λ)

The method depends on counting the α-particles emitted per second by a known quantity of the substance. Each disintegrating atom emits only one α-particle. Hence the number of α-particles counted must be equal to the number of atoms disintegrating.

From the general relation,

$$-\frac{dN}{dt} = \lambda N$$

we have $\lambda = \dfrac{dN/dt}{N}$

The decay constant λ is equal to the ratio of the atoms disintegrating per unit time to the total number of atoms present.

Geiger and Rutherford found by the method of scintillations that one gramme of thorium, which contains 2.61×10^{21} atoms, emits 4500 α-particles per second.

$$\therefore \quad \lambda = \frac{4500}{2.61\times10^{21}} = 1.72 \times 10^{-18}\, s^{-1} = 5.44 \times 10^{-11} y^{-1}$$

$$\therefore \quad T = \frac{0.693}{\lambda} = \frac{0.693}{5.44\times10^{-11} y^{-1}} = 1.28 \times 10^{10}\ \text{yrs.}$$

16.12 Applications of radioactivity

1. Nuclear radiations like γ-rays have been utilized for the preservation of food. Food-stuff mainly meat, poultry, fish, fruits etc are exposed to γ-rays from cobalt-60 or caesium-137. A dose of about 2 to 5 million rads is sufficient to destroy almost all bacteria in food. This increases the shelf-life of the food-items without refrigeration. Radiation is also being used for insect disinfection of wheat and flour.

2. Gamma radiation from cobalt-60 is used in hospitals to sterilize materials like hypodermic syringe, surgical instruments, dressings, etc.

3. Radiation can also be used as pesticide. Population of insects which causes considerable damage to both plant crops and livestock can be controlled by irradiating the male members of these insects which render them sterile.

4. Radiation is also used for producing new and improved varieties of plants. This is achieved by plant mutations produced by radiation. Radioisotopes can also be used to raise crop-yields.

5. Gamma radiations from radioactive sources like cobalt-60, iridium-192 are used in industrial radiography *i.e.*, for investigating the interiors of metallic castings for detecting any flaws or defects.

6. A carefully prepared mixture of radio-thorium (α-emitter) with zinc sulphide exhibits a more or less permanent luminescence and is used for coating the pointers and figures of clocks and watches, for rendering visible signs in theatres and so on.

7. Medical applications: Radioisotopes are used to diagnose the nature of blood circulatory disorders, defects of bone metabolism, to locate tumors, etc. Radio-sodium is used to study the circulatory disorder in blood vessels while radioactive iodine is used to study any disorder in thyroid gland. Tc^{99m} is used to study the functioning of different organs like liver, kidney and spleen under normal and diseased conditions. Radioisotopes are used to detect and locate the presence of tumours – particularly brain tumours which are difficult to detect and specially to locate.

Radiations from radium, cobalt-60 are used in radiotherapy – treatment of cancer by radiation.

16.13 Radioactive Dating: The Age of the Earth

The age of the earth is estimated from the relative abundance of the two isotopes of uranium, U^{238} *and* U^{235}. The present abundance ratio of U^{235} and U^{238}, is 1.140 (0.7% to 99.3%). The half-lives of U^{235} and U^{238}, according to the best estimate are 7.07×10^8 years and 4.5×10^9 years respectively. Assuming that at the beginning the proportions of the two isotopes were equal, the present relative abundance of U^{238} and U^{235} may be expressed as

$$\frac{N_1}{N_2} = \frac{99.3}{0.7} = \frac{N_0 e^{-\lambda_1 t}}{N_0 e^{-\lambda_2 t}} = e^{(\lambda_2 - \lambda_1)t}$$

$$\text{where } \lambda_1 = \frac{0.693}{4.5\times10^9} y^{-1} \text{ and } \lambda_2 = \frac{0.693}{7.07\times10^8} y^{-1}$$

$$\therefore \quad \ln\left(\frac{99.3}{0.7}\right) = (\lambda_2 - \lambda_1)\, t$$

$$\text{or} \quad t = \frac{1}{\lambda_2 - \lambda_1} \ln\left(\frac{99.3}{0.7}\right)$$

$$\cong 5.93 \times 10^9 \text{ years}$$

$$\cong 5{,}000 \text{ million years.}$$

This value agrees nearly with that given by astronomical evidence for the age of the universe.

Radioactive Carbon Dating

All plants use CO_2 from the atmosphere for growth. A portion of carbon is C^{14} which is radioactive. So all plants are slightly radioactive. When a plant dies, it does not take CO_2 from the atmosphere any more and so no additional C^{14} is taken in. C^{14}, already present within the plant body begins to decay without being replaced. Measurement of the relative amount of C^{14} in an organic archeological sample provides a sensitive method of dating.

16.14 Biological effect of radiation

When radiation like α-ray, β-ray, γ-ray or neutrons pass through a medium, they cause ionization of the atoms of the medium either directly or indirectly. The ion-pairs (electrons and positive ions) thus produced dissipate their energy by ionizing nearby atoms to produce secondary ion-pairs. The total number of ion-pairs produced, both primary and secondary, is proportional to the energy that the radiation deposits in the medium. When the medium is the human body, the radiation interacts with the atoms of the living cells and either ionizes or excites the atoms of the living cells. As a result some of the cell constituents are altered or destroyed by ionization. The radiation also interacts with water, the major constituent of human body, and produces hydrogen peroxide. Hydrogen peroxide is a strong oxidizing agent and also damages the cells. Examples of damages are the breaking up of chromosomes, swelling of the nucleus of a cell or of the entire cell and changes in the permeability of the cell membrane. As a result the ability of the cell to divide may be destroyed leading to the eventual death of the cell.

The biological effects can be divided into three groups:

(i) Short-term recoverable effects;

(ii) Long-term irrecoverable effects; and

(iii) The genetic effect.

The first two types of effects are limited to the individuals actually receiving the radiation and is referred to as *somatic* effects. The *genetic* effects, on the other hand, are not evident in the irradiated person but appear in their later generations.

The nature and extent of damages depends on the type of radiation, on the depth to which radiation has penetrated, on the extent of the body exposed, on the amount of radiation absorbed and also upon whether the exposure was *chronic i.e.*, exposure received in small amount but repeated over a prolonged period or *acute i.e.*, the exposure received in one large dose.

Acute effects

Acute whole-body radiation exposure affects all the organs and systems of the body. However, since not all the organs or organ systems are equally sensitive to radiation, the pattern of response, or disease syndrome, in an overexposed individual depends on the magnitude of the dose. In order of increasing severity these syndromes are (i) the *hemopoietic syndrome*, (ii) the *gastrointestinal syndrome*, and (iii) the *central nervous system syndromes*. Certain effects are common to all these categories. The effects are

Nausea and vomiting

Malaise and fatigue

Increased temperature

Blood changes

When excessive doses are absorbed, the first noticeable effects are on the blood-forming tissues. A drop in the white blood cell count becomes evident in the first few hours after the exposure. This is followed by a sickness pattern of diarrhoea, vomiting and fever which is now called radiation sickness. Recovery is possible from small acute doses. Large doses are lethal within a few weeks. Smaller doses produces short term effects, such as skin disorders and loss of hair, which are generally recoverable. More serious is the damage done to the bone-marrow and to the other cells which may not have the ability to repair the damage. This leads to leukemia and to the production of cancerous cells

and finally to malignant tumours. When the dose is in excess of 20 Gy (2000 rads), the central nervous system as well as all the other organ system in the body are damaged. Unconsciousness follows within minutes after exposure and death in a matter of hours to several days. The rapidity of onset of unconsciousness is directly related to dose.

Exposure to a single large exposure or continuing (chronic) low-level exposure may lead to delayed effects which take about 5 to 20 years. Continuing over-exposure may be due to exposure to external radiation fields or can result from inhalation or ingestion of a radioisotope which then becomes fixed in the body through chemical reaction with the tissue protein or, because of the chemical similarity of the radioisotope with normal metabolites, may be systematically absorbed within certain organs and tissues. In either case, the internally deposited radioisotope may continue to irradiate the tissue for a long time. Among the delayed consequences of over exposure that are of concern are cancer, genetic effects, shortening of life span, and cataracts.

16.15 Radiation Dosimetry

Radiation damage (*i.e.*, chemical and biologic changes in tissue) depends on the absorption of energy from the radiation rather than upon the amount of ionization that the radiation produces in air and is approximately proportional to the concentration of absorbed energy in tissue. For this reason, the basic unit of radiation dose is expressed in terms of absorbed energy per unit mass of tissue. This unit is called the *gray* (Gy).

Gray

One gray is an absorbed radiation dose of 1 joule per kilogram.

$$1 \text{ Gy} = 1 \text{ J/kg}$$

The gray is the SI unit of dose and is universally applied to all types of ionizing radiation.

Rad

Before the introduction of the SI units, radiation dose was measured by a unit called the *rad* (radiation absorbed dose).

One rad is an absorbed dose of 100 ergs per gram.

1 rad = 100 ergs/gm

1 Gy = 1 J/kg = 10^7 ergs/1000gm = 10^4 ergs/gm = 100 rads.

Exposure Unit

Radiation fields to which an organism may be exposed are frequently specified in **exposure units.** One exposure unit is defined as that quantity of radiation (X-ray or γ-ray) that produces, in air, ions carrying 1C of charge (of either sign) per kilogram of air.

1 X unit = 1 C/kg of air.

X unit is the SI unit of exposure. Before the introduction of the SI unit, the unit of X-ray exposure was called the *roentgen* and was symbolized by R. The roentgen was defined as that quantity of X-or gamma radiation that produces 1sC of charge of either sign per cubic centimeter of air at 0° and 760 mm Hg.

Since 1 ion carries a charge of 4.8×10^{-10} sC, and the mass of $1cm^3$ of standard air is 0.001293 gm, it can be shown that 1 roentgen is equivalent to

$$1 \text{ R} = 2.58 \times 10^{-4} \text{ C/kg of air.}$$

It can be further shown that an exposure of 1R corresponds to an absorption of 87.7 ergs/gm of air, or to a dose to the air of 0.877 rad.

The Rem

Usually, chemical and biologic effects of radiation depend not only on the amount of energy absorbed in the irradiated medium but also upon the distribution of the absorbed energy within the medium. For equal absorbed doses, various types of ionizing radiation often differ in the efficiency with which they elicit a particular chemical or biologic change. The efficiency with which a particular type of radiation evokes a particular chemical or biologic change is described by a term called *relative biologic effectiveness (RBE)*. The RBE of a radiation is defined as the ratio of the amount of energy from a reference radiation – usually 200 keV X-rays or

C_0^{60} radiation, to produce a given biological effect to the amount of energy from that radiation to produce the same effect.

$$\text{RBE} = \frac{\text{Dose of reference radiation required to produce a particular response}}{\text{Dose of radiation in question required to produce the same response}}$$

The *RBE* dose in rem (acronym for roentgen equivalent man) is the product of the RBE and the dose in rads.

RBE dose (rem) = Absorbed dose (rad) × RBE.

Often, the effectiveness with which different types of radiation produce a particular biological effect varies with the linear energy transfer (LET) of the radiation. *The linear energy transfer (LET)* is the average loss in energy per unit length of path of the incident radiation and is the product of the *specific ionization* produced by the radiation and W-quantity. Specific ionization is the number of primary and secondary ion pairs produced per unit length of path of the incident radiation while W-quantity is the average energy expended in creating an ion-pair produced in air and is equal to 33.7 eV.

The *Dose-equivalent (DE)* is the product of the dose in rads and a *quality factor (QF)* which varies with the LET of the radiation.

DE (rem) = D (rad) × QF.

The quality factor of some radiations is given in the table below.

Type of radiation	QF
X-rays, γ-rays and β-particles	1.0
Neutrons and protons ≤ 10MeV	10 (30 for irradiation of eye)
α-particles from natural radionuclide	10
Heavy recoil nuclei	20

The SI unit of dose equivalent is *Sievert.*

DE (Sievert) = Dose (Gray) × QF.

Example 16.8 *Some amount of a radio-active substance of half-life 30 days is spread inside a room. Consequently the level of radiation inside the room became 50 times the permissible level for normal occupancy of the room. After how many days the room would be safe for occupation?*

Soln.

$$N = N_0 e^{-\lambda t} \qquad \frac{N}{N_0} = \frac{1}{50}$$

$$\frac{N}{N_0} = e^{-\lambda t} \qquad \lambda = \frac{0.693}{30} d^{-1} = 0.0231 d^{-1}$$

or

$$e^{\lambda t} = \frac{N_0}{N} = 50$$

$$\lambda t = \ln(50) = 3.912$$

$$t = \frac{3.912}{0.0231} = 169.35 \text{ days}$$

Example 16.9 *A sample of pitchblende has a lead-uranium weight ratio of 9/40. Calculate the age of the mineral. (Given: half-life of uranium = 4.1×10^9 y and the atomic weights of lead and uranium are 206 and 238.4 respectively. Assume that lead of non-radioactive origin is absent).*

Soln.

Weight ratio of lead-uranium = $\frac{9}{40}$. This means that if there is 9 gms of lead, then the weight of uranium present is 40gms.

No. of lead atoms in 9 gm of lead

$$= \left(\frac{9}{206}\right) \times (6 \times 10^{23}) = 0.04369 \times (6 \times 10^{23}) \text{ atoms}$$

No. of uranium atoms in 40 gm of uranium

$$= \left(\frac{40}{238.4}\right) \times (6 \times 10^{23}) = (0.16779) \times (6 \times 10^{23}) \textit{ atoms}$$

Total no. of uranium atoms in the beginning

$$N_0 = \left(\frac{9}{206} + \frac{40}{238.4}\right) 6 \times 10^{23} \text{ atoms}$$

$$= (0.04369 + 0.16779)\ 6 \times 10^{23} \text{ atoms}$$

$$= 0.21047 \times 6 \times 10^{23} \text{ atoms.}$$

From the relation

$N = N_0 e^{-\lambda t}$; we have

$$\frac{N}{N_0} = e^{-\lambda t}$$

Now $\frac{N}{N_0} = \frac{(0.16779) \times 6 \times 10^{23}}{(0.21047) \times 6 \times 10^{23}} = \frac{0.16779}{0.21047}$

or $e^{\lambda t} = \frac{N}{N_0} = \frac{0.21047}{0.16779}$

$= 1.25436.$

$$\lambda = \frac{0.693}{4.1 \times 10^{9}} y^{-1} = 1.69 \times 10^{-10} y^{-1}.$$

$\lambda t = \ln (1.25436) = 0.22663$

$$t = \frac{0.22663}{1.69 \times 10^{-10}} y$$

$= 1.34 \times 10^9$ y.

***Example 16.10** A piece of an ancient wood boat shows an activity of C^{14} of 3.9 disintegrations per minute per gram of carbon. Estimate the age of the boat if the half-life of C^{14} is 5568 years. Assume that the activity of fresh C^{14} is 15.6 disintegrations per minute per gram.*

Soln.

Let the age of the boat be t years.

From $N = N_0 e^{-\lambda t}$, we have

$$\frac{N}{N_0} = e^{-\lambda t}$$

Here

activity $\lambda N = 3.9$

or $$e^{\lambda t} = \frac{N_0}{N} = \frac{15.6}{3.9}$$

and $\lambda N_0 = 15.6$

$$\therefore \quad \frac{\lambda N_0}{\lambda N} = \frac{N_0}{N} = \frac{15.6}{3.9}$$

or $$\lambda t = \ln\left(\frac{15.6}{3.9}\right)$$

$$\lambda = \frac{0.693}{5568} y^{-1} = 1.24 \times 10^{-4}\ y^{-1}.$$

or $$t = \frac{\ln\left(\frac{15.6}{3.9}\right)}{\lambda}$$

$$= \frac{\ln\left(\frac{15.6}{3.9}\right)}{1.24\times10^{-4}\ y^{-1}} = 1.118 \times 10^{4} \text{ yrs.}$$

***Example 16.11** Calculate the activity of K^{40} in 100 kg man assuming that 0.35% of the body weight is potassium. The abundance of K^{40} is 0.012%; its half-life is 1.31×10^{9} yrs.*

Soln.

Total mass of potassium in 100 kg man

$$= \frac{0.35}{100}\times100\text{kg} = 0.35\text{kg}.$$

Therefore, the amount of K^{40} is

$$\frac{0.012}{100}\times0.35 = 4.20\times10^{-5}\text{ kg}.$$

Number of atoms in 4.20×10^{-5} kg of K^{40}

$$= \frac{6.023\times10^{23}\times10^{3}}{40}\times4.20\times10^{-5}$$

$$= 6.32425\times10^{20} \text{ atoms}$$

Activity of K^{40} $$\lambda = \frac{0.693}{1.31\times10^{9}\times365\times24\times60\times60}$$

$$= \lambda N$$

$$= \frac{0.693 \times 6.32425 \times 10^{20}}{1.31 \times 10^9 \times 365 \times 24 \times 60 \times 60}$$

$= 1.061 \times 10^4$ disintegrations per second

$= 0.287\ \mu$Ci.

***Example 16.12** Radon, the disintegration product of radium is in equilibrium with 1 gram of radium. Find the mass of radon. Half-life of Ra^{226}=1590 years, Rn^{222}=3.82 days.*

Soln.

For equilibrium

$$\lambda_1 N_1 = \lambda_2 N_2$$

$$\left(\frac{0.693}{1590}\right) \times 1 = \left(\frac{0.693 \times 365}{3.82}\right) \times N_2$$

$$\therefore \quad N_2 = \frac{3.82}{1590 \times 365}$$

$$= 6.5 \times 10^{-6} \text{ gm.}$$

Here

$$\lambda_1 = \frac{0.693}{1590} y^{-1}$$

$$\lambda_2 = \frac{0.693}{3.82} d^{-1} = \frac{0.693 \times 365}{3.82} y^{-1}$$

$N_1 = 1$ gm, $\quad N_2 = ?$

***Example 16.13** Calculate the number of α-particles emitted per second from one gram of radium. The atomic weight of radium is 226 and its half-life is 1600 years. Calculate the volume of helium gas produced by 1gm of radium in one year.*

Soln.

$$\left(-\frac{dN}{dt}\right) = \lambda N$$

$$= \left(\frac{0.693}{1600 \times 365 \times 3600}\right) \times \left(\frac{6.023 \times 10^{23}}{226}\right)$$

$= 3.66 \times 10^{10}$ per second.

Here

$$\lambda = \frac{0.693}{1600 \times 365 \times 3600} s^{-1}$$

$$N = \frac{6.023 \times 10^{23} \times 1}{226}$$

Therefore, the number of α-particles emitted from one gram of radium in one second$=3.66\times10^{10}$

$\therefore$ Number of α-particles emitted in one year

$$= 3.66 \times 10^{10} \times 365 \times 24 \times 3600$$

$$= 1.154 \times 10^{18}$$

Volume of α-particles at NTP

$$V = \frac{(22.4)\times(1.15)\times10^{18}}{6.023\times10^{23}} \text{ litres}$$

$$= 4.29 \times 10^{-5} \text{ litres.}$$

EXERCISE

1. What is radioactivity? Does it depend on any external agent? Explain the displacement law of radioactivity as put forward by Soddy and Fajans.

2. Mention the properties of α, β and γ-rays. Describe Curie's experiment to show the constituent radiations.

3. Derive an expression for the law governing radioactive decay.

4. What do you mean by half-life of a radioactive substance? Hence obtain an expression for the half-life.

5. What is mean life of a radioactive element? Obtain an expression for the mean life of a radioactive substance.

6. Explain the terms decay constant, half-life and average life as applied to a radioactive substance. Find the relations between them.

7. What are the units of radioactivity? Find out the mass in gm of radium – B ($_{82}Pb^{214}$) that will give 1 curie and 1 rutherford of activity.

8. Explain the terms isotopes, isobars, isotones and isomers. Give examples.

9. Give the theory of successive disintegration of radioactive substances.

10. Explain what is radioactive equilibrium. Distinguish between secular and transient equilibria.

11. What is radioactive decay constant? How would you measure the decay constants of elements of short half-life and long half-life?

12. Discuss the different applications of radioactivity. Explain what is meant by radioactive dating?

13. Discuss briefly the biological effects of radiation. Explain the terms roentgen, rad, gray, relative biological effectiveness (RBE) and quality factor with this connection.

14. Assuming that it were to remain undisturbed since 1898 A.D. how much of Madam Curie's 200 mg of radium would be left in the year 8378 A.D. (Half-life of radium is 1620 years) [12.5 mg]

15. A certain radioactive substance has a disintegration constant $\lambda=1.44\times10^{-3}$ per hour. In what time will 75% of the initial number of atoms disintegrate? [962.9 hours]

16. 1 gram of radioactive substance disintegrates at the rate of 3.7×10^{10} disintegration per second. The atomic weight of the substance is 226. Calculate its mean life. [2262 years]

17. Calculate the time required for 10% of a sample of thorium to disintegrate. Assume the half-life thorium to be 1.4×10^{10} years. [2.1×10^{9} years]

18. A radio-active element due to accident in the research laboratory gets embedded in its floor and walls. The initial rate of disintegration is 64 times the safe limit. The half-life of the radioactive material is 32 days. Calculate the time after which the laboratory will be ready for safe use. [192 days]

19. A carbon specimen found in a cave contained 1/8 as much C^{14} as an equal amount of carbon in living matter. Calculate the approximate age of the specimen. Half-life of C^{14} is 5568 years. [≅ 16710 years]

20. C^{14} has an abundance of 1.3×10^{-12} in natural carbon and has a half-life of 5730 years. Calculate the number of disintegrations per hour in one gram of natural carbon. [771.8]

21. The number of α-particles emitted per gram of radium each second is 3.71×10^{10}. Atomic weight of radium is 226. Calculate its half-life in years. [1.5791×10^{3} years]

22. How much time will it take for a 8 mCi source to reduce to 1 mCi source? Half-life of the source =10 years. [30 years]

23. The wood of a living tree contain C^{14} activity of 12 disintegrations per minute per gram. During an archeological survey a piece of wood from the ruins of an ancient dwelling was found to have a mass of 36 grams and C^{14} activity of 224 disintegrations per minute. Estimate the age of the archeological find. Half-life of C^{14} is 5568 years. [5277 years]

24. A piece of wood weighs 50 grams and shows C^{14} activity of 320 disintegrations per minute. Estimate the length of time which has elapsed since this piece of wood was part of a living tree. Assume that a living plant shows a C^{14} activity of 12 disintegrations per minute per gram. Half-life of C^{14} is 5730 years. [5197.5 years]

25. What is the mass of one curie and one rutherford of U^{234}? [1.438×10^{-11} gm; 3.887×10^{-16} gm]

CHAPTER XVII

STATISTICAL MECHANICS

Many physics experiments may be analyzed as if the interactions take place in single, isolated events. For example, in the emission of light from the atoms of a gas at low density, the transitions of an electron in any atom are unlikely to be affected by the presence of other atoms. Thus we can treat the light from the collection of many atoms as we would treat light from a single atom. On the other hand, let us consider the addition of energy to a gas in a container by raising its temperature. If a total energy E is added to a gas of N atoms, it cannot be predicted with certainty how much energy any particular atom will acquire. On the average, the energy of each atom will increase by E/N, but some atoms might acquire no additional energy at all, while others might absorb 10 E/N or even 100 E/N.

The sharing of energy among the many parts of a system cannot be simply analyzed in terms of single isolated events. The analysis of such *cooperative* phenomena requires the techniques of *statistical physics*, in which we are not concerned with calculating the exact outcome of single isolated events, but rather predicting the average outcome of many cooperative events, based on the *statistical distribution* of the possible outcomes.

Statistical mechanics is that branch of theoretical physics in which the microscopic properties of a system are studied on the basis of molecular kinetic concepts and the methods of mathematical statistics. It deals with systems that are in an equilibrium state or near to one. Statistical mechanics is not concerned with the actual motions or interactions of individual particles, but rather with their most *probable* behaviour. It can be applied with equal facility to classical systems (notably molecules in a gas) and to quantum mechanical systems (notably photons in a cavity and free electrons in a metal).

17.1 The macroscopic and microscopic states

Let us consider a physical system composed of N identical particles confined to a space of volume V. In the so called *thermodynamic limit* $N \to \infty$ as $V \to \infty$ so that the ratio N/V, called

the particle density and denoted by the symbol n, stays fixed at a preassigned value. In this limit, the *extensive* properties of the system become directly proportional to the *size* of the system (*i.e.*, proportional to N or to V), while the *intensive* properties become independent thereof. The particle density, of course , remains an important parameter for *all* physical properties of the system.

If the particles comprising the system are noninteracting, the total energy E of the system will be equal to the sum of the energies ε_i of the individual particles:

$$E = \sum_i n_i \varepsilon_i \tag{17.1}$$

where n_i denotes the number of particles with energy ε_i. Clearly

$$N = \sum_i n_i \tag{17.2}$$

According to quantum mechanics, the possible values of the single particle energies ε_i are discrete and their magnitude depends crucially on the volume V to which the particles are confined. Accordingly, the possible values of the total energy E are also discrete. However, for large V, the spacings of the difference between the different energy values are so small in comparison of the total energy of the system that the parameter E might be regarded as almost *continuous variable*. This will be true even if the particles were mutually interacting. However, then the total energy will not be given by eqn. (17.1).

The specification of the actual values of the parameters N, V, E then defines a particular *macrostate* of the given system.

In general, there will be a large number of different ways in which a macrostate (N, V, E) can be specified. If the system is non-interacting, then E will, simply be sum of the N single particle energies ε_i. Obviously there will be a large number of different ways in which the individual ε_i's can be chosen so as to make the total energy equal to E. In other words, there will be a large number of different ways in which the total energy E of the system can be distributed among the N particles constituting it. Each of these (different) ways specifies a particular *microstate* or *complexion*, of

the given system. A given macrostate of the system, in general, correspond to a large number of microstaes. It is only natural to assume that at any time the system is equally likely to be in any one of these microstates. This assumption forms the backbone of our formalism and is generally referred to as the postulate of "*equal a priori probabilities*" for all microstates consistent with a given macrostate.

This is illustrated with a couple of examples.

Let there be cell 1, cell 2, cell 3, ..., cell i in phase space. Suppose there are four phase points abcd in cell 1, three phase points efg in cell 2, one phase point h in cell 3 and two phase points jk in cell i as shown below (Fig. 17.1)

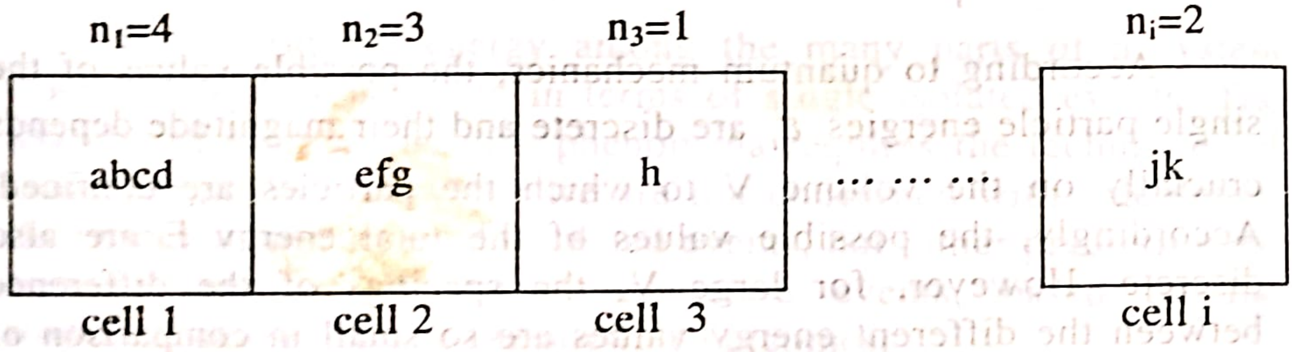


Fig. 17.1

The macrostate in Fig. 17.1 is specified merely by giving the phase points $n_1=4$, $n_2=3$, $n_3=1$, ... $n_i=2$ of different cells. This also represents a particular microstate by specifying the position of phase points abcd in cell 1, efg in cell 2 and so on. Now if the phase point a and e from different cells are interchanged, then the microstate is changed because the positions of the two phase points are changed. But the macrostate remains unchanged as the number of phase points in each cell remains the same. In a similar manner different microstates can be considered in the same macrostate by interchanging the positions of the phase points. Thus different microstates may correspond to the same macrostate.

Let us roll two dice P and Q. Any of the six sides may turn up; any face is just as likely to turn up as another. Let the number turning up on P be denoted by P_1, P_2, P_3, ...P_6 while those on Q by Q_1, Q_2, Q_3, ...Q_6. If the two dice are rolled together, the *sum of the number* (S) turning up on the two dice may be from 2 to 12. In all there are 36 possible combinations. If we want a particular number *i.e.*, a particular pattern or macrostate to turn up, there may be one or more ways. If we want 12 to turn up, there is only one way P_6Q_6. If we want 8 to turn up, then the chances are P_2Q_6, P_6Q_2, P_5Q_3, P_3Q_5, P_4Q_4. That is, there are five microstates by which the macrostate 8 may exist. One microstate is as possible as other. For macrostate 12, there is only one microstate, whereas, for 8 there are five. Thus, the macrostate which would be most observed is the most probable state.

Thermodynamic quantities describing the macroscopic states of systems involve

i. mechanical quantities such as pressure and energy.

ii. non-mechanical quantities such as temperature and entropy and

iii. external parameters such as volume, intensity of the field force due to neighbouring systems (bodies) etc.

The connection between statistical and thermodynamic conceptions of a macroscopic assembly is established through the consideration of the temperature and entropy of the assembly.

Phase Space

To specify the position as well as energy of a molecule inside a gas, we must specify three space coordinates x,y,z and three momentum coordinates p_x, p_y, p_z. From purely mathematical concept, a six dimensional space may be imagined where we have the six coordinates x,y,z, p_x, p_y, p_z. This six dimensional space for a single molecule is called *phase space*. The instantaneous state of a particle in the phase space is represented by a point known as *phase point*. A six dimensional element of volume dx, dy, dz, dp_x, dp_y, dp_z, called a *cell*, may be considered in the phase space. Thus the phase space may be divided into a large number of cells, each cell in its turn may contain a large number of phase points.

Ensembles

A system is defined as a collection of a number of particles. *An ensemble is defined as a collection of a large number of macroscopically identical, but essentially independent systems.* By *macroscopically independent* is meant that each of the systems constituting the ensemble satisfies the same macroscopic conditions, *e.g.*, volume, energy, pressure, total number of particles, etc. The term *independent system* means that the systems constituting the ensemble are mutually *non-interacting*. In an ensemble the systems play the same role as the non-interacting molecules do in a gas.

Probability

The probability of an event may be defined as the ratio of the number of cases in which the event occurs to the total number of cases. That is,

$$\textit{the probability of an event} = \frac{\text{total of cases in which the event occurs}}{\text{total numer of cases}}$$

Suppose, a coin is tossed a large number of times and the number of times the *head* or *tail* is uppermost is counted. Experience show that the heads and tails will be uppermost equal number of times. That is

$$P(\text{head}) = \frac{1}{2} \quad \text{and} \quad P(\text{tail}) = \frac{1}{2}.$$

Thermodynamic Probability

The number of microstates corresponding to a given macrostate is called its thermodynamic probability usually denoted by W.

Let us consider two cells in phase space represented by i and j and four molecules a,b,c and d. If N_i and N_j be the number of molecules in cells i and j respectively, then the possible macrostates will be five in number as shown below.

N_i	4	3	2	1	0
N_j	0	1	2	3	4

In general, different number of microstates will correspond to each of the five macrostates. Let us consider the microstates corresponding to the macrostate $N_i=3$, $N_j=1$

Cell i	abc	abd	acd	bcd
Cell j	d	c	b	a

Thus there will be 4 microstates corresponding to the macrostae $N_i = 3$, $N_j = 1$. Hence the thermodynamic probability for this macrostates is 4 or $W = 4$.

17.2 Fundamental Postulates of Statistical Mechanics

In statistical mechanics, we assume certain fundamental postulates. These are

i. Any gas consists of a large number of molecules which are always in motion and behave like very small elastic spheres.

ii. All the cells in the phase space are of equal size.

iii. All accessible microstates corresponding to possible macrostates are equally probable. This is called the *postulate of equal a priori probability.*

iv. The equilibrium state of a gas corresponds to the macrostates of maximum probability.

v. The total number of molecules is constant.

vi. The total energy of the system is constant.

17.3 Boltzmann's theorem on entropy and probability

Boltzmann established a relation between probability (a statistical quantity) and entropy (a thermodynamical quantity). Boltzmann argued that the equilibrium state of a system is the state of maximum probability, *i.e.*, the probability of the system in equilibrium is maximum. But from the thermodynamic point of

view, the equilibrium state of the system is the state of maximum entropy. If the system is not in equilibrium, then changes take place within the system until the equilibrium state is reached *i.e.*, the state of maximum entropy is reached. Thus in the equilibrium state, both the thermodynamic probability and entropy have their maximum values. Since an equilibrium state is the state of maximum entropy and maximum probability, Boltzmann argued that *entropy must be a function of probability*.

That is,

$$S = f(W) \qquad \text{(i)}$$

where S is the entropy and W is the thermodynamic probability of the state.

Let S_1 and S_2 be the entropies and W_1 and W_2 be the thermodynamic probability of two separate systems.

Then $S_1 = f(W_1)$ and $S_2 = f(W_2)$ (ii)

The total entropy of the two system is

$$S_1 + S_2 = f(W_1) + f(W_2) \qquad \text{(iii)}$$

But the thermodynamic probability of the two systems taken together is W_1W_2.

$$\therefore \quad f(W_1W_2) = f(W_1) + f(W_2) = S_1 + S_2 \qquad \text{(iv)}$$

For eqn. (iv) to be valid, f(W) must be a logarithmic function of W.

$$\therefore \quad f(W) = k \log W \quad \text{or} \quad S = k \log W \qquad (17.3)$$

17.4 Statistical equilibrium

To determine the equilibrium state of a system, Boltzmann applied the *canonical principle* according to which, *the equilibrium state of a system is that which is most probable.*

Let us consider an isolated thermodynamical system which consists of N particles. The energy states available to the particles are E_1, E_2, E_3..... etc. The energy states which may be continuous or quantized are due to vibrational and rotational energy of the particles. At a particular time the particles are distributed among the

different energy states, so that n_1 particles have energy E_1; n_2 particles have energy E_2 and so on. The total number of particles is

$$N = n_1 + n_2 + n_3 + ... = \sum_i n_i \qquad \text{(i)}$$

Here i = 1, 2, 3, ...

Assuming that the total number of particles remains constant for all processes occurring in the system, the total energy of the system is

$$E = n_1E_1 + n_2E_2 + n_3E_3 + ... = \sum_i n_iE_i \qquad \text{(ii)}$$

If the system is isolated, the total energy E must be constant, *i.e.*,

$$E = \sum_i n_iE_i = \text{constant.} \qquad (17.4)$$

Consider a gas having N molecules at a certain temperature and pressure. If the system is isolated *i.e.*, its pressure, temperature and volume are kept constant, then the total energy of the system must remain constant. But, the molecules of the gas collide with each other and also with the walls of the containing vessel. Suppose a collision takes place between a fast molecule and a slow molecule. After the collision the fast molecules may have slowed down while the slow ones may have sped up. Hence the particles after the collision are in different energy states. In other words, the numbers $n_1, n_2, n_3, ...$, which give the distribution of the N particles among the available energy states, may be changing. It can be reasonably assumed that for each microscopic state of a system of particles, there is a distribution which is more favoured than any other. In other words, we may say that, *given the physical conditions of the system of particles, there is a most probable distribution. When this distribution is achieved, the system is said to be in statistical equilibrium*. Hence, the basic problem in statistical thermodynamics is to obtain the most probable distribution law for a given composition of the system.

The three most probable distribution laws used in practice are

(i) Maxwell-Boltzmann Distribution Law

(ii) Fermi-Dirac Distribution Law (Statistics)

(iii) Bose-Einstein Distribution Law (Statistics).

17.5 Maxwell Boltzmann Distribution Law

Consider a system that contains a large number of *identical* but *distinguishable* particles of any spin. The molecules of a gas are the particles of this kind. This kind of particles obey the Maxwell-Boltzmann Distribution Law. The law tells us how a given amount of energy is distributed among the various members of an assembly of identical particles in the most probable distribution. Let us derive the M-B distribution law.

Consider that the system contains N number of identical but distinguishable particles. Identical because the particles have the same structure and distinguishable because the particles may be in different energy states at a given instant. As shown in the figure below (Fig. 17.2),

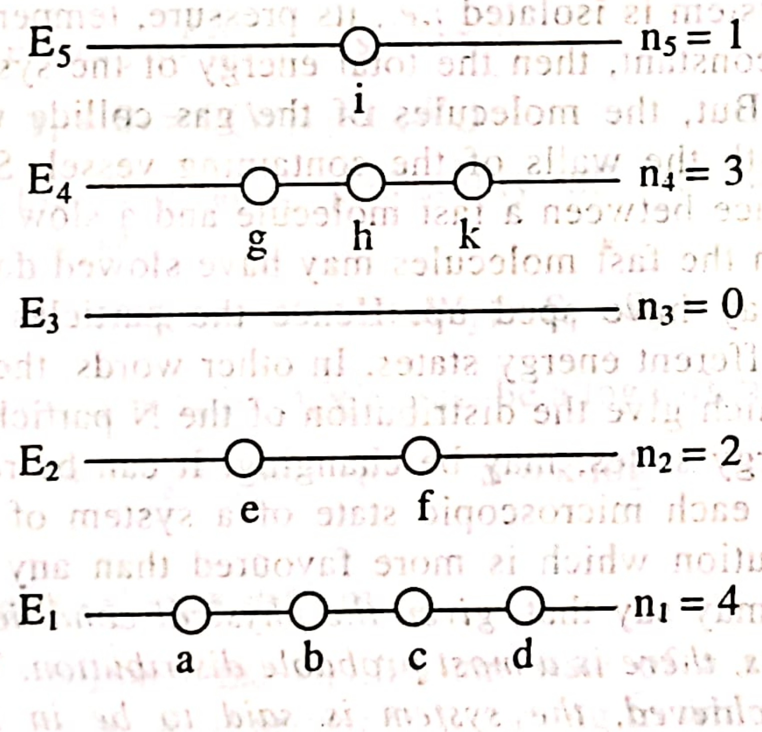


Fig. 17.2

4 particles are in energy state E_1, 2 in state E_2, zero in state E_3 and so on. It is assumed that all the energy states are accessible to each particle. Consequently, it can be assumed that the probability of any particular partition is proportional to the number of different ways in which the particles can be distributed in the existing available energy states so as to produce the desired partition.

To obtain the distribution as shown in Fig. 17.2, the first particle *a* in the energy state E_1 can be selected in N different ways. There are now (N-1) particles left. So the second particle *b* in state E_1 can be selected in (N-1) different ways and so on. Therefore, the total number of ways in which the first four particles in state E_1 can be selected is given by

$$N(N-1)(N-2)(N-3)=\frac{N!}{(N-4)!}$$

Moreover, the four particles in state E_1 can be arranged in 4! different orders, *e.g.*, *abcd, bcda, cdab* and so on. There are 24 different ways. But, it is immaterial for these particles to be arranged in any particular order in state E_1, because they are identical. Thus, the total number of distinguishable different ways are,

$$\frac{N!}{4!(N-4)!}$$

In general, if the first state E_1 consists of n_1 particles, the distinguishable different ways in which n_1 particles can be arranged in state E_1 are,

$$p_1=\frac{N!}{n_1!(N-n_1)!} \qquad \text{(i)}$$

For the second state E_2, the number of available particles is $(N-n_1)$. If the number of distinguishable particles in state E_2 is n_2, then the number of distinguishable ways in which these n_2 particles may be arranged are,

$$p_2=\frac{(N-n)!}{n_2!(N-n_1-n_2)!} \qquad \text{(ii)}$$

If this process is continued for all the available states, the total number of distinguishable ways are obtained by multiplying p_1, p_2, p_3, etc. Or,

$$P = p_1\times p_2\times p_3\times ...$$

$$=\left(\frac{N!}{n_1!(N-n_1)!}\right)\times\left(\frac{(N-n_1!)}{n_2!(N-n_1-n_2)!}\right)\times ...$$

$$= \frac{N!}{n_1!n_2!n_3!...}$$

As shown in Fig. 17.2, the distinguishable ways are,

$$P = \frac{N!}{4!3!\,2!\,1!...} \qquad \text{(iv)}$$

Here 0! is equal to 1.

It was assumed so far that all the available states have the same probability of occupation by the particles. However, it may happen that a particular energy state may be favourable with more different angular momentum states than the rest and hence it is more likely to be occupied. In other words, different energy states may have different intrinsic probabilities, say g_i. If this intrinsic probability factor is taken into account, the value of *P* will be different.

Let g_i be the probability of locating a particle in a certain energy states E_i. Then the probability of locating 2 particles in the same energy state is $g_i \times g_i = g_i^2$. For *n* particles the probability will be $g_i^{n_i}$. Hence the total probability for a given distribution is given by

$$P = \frac{N!g_1^{n_1}\,g_2^{n_2}\,g_3^{n_3}\,...}{n_1!\,n_2!\,n_3!...} \qquad \text{(v)}$$

Here n_1, n_2, n_3... are the number of particles in the states E_1, E_2, E_3... and g_1, g_2, g_3... are the intrinsic probabilities for the states E_1, E_2, E_3...

If the particles are further assumed to be indistinguishable, *i.e.*, particles in state E_1 cannot be distinguished from particles in state E_4, then N! permutations among the particles themselves and occupying the different states result in the same distribution. The probability in this case is given by

$$P = \frac{1}{N!}\left[\frac{N!g_1^{n_1}\,g_2^{n_2}\,g_3^{n_3}\,...}{n_1!\,n_2!\,n_3!...}\right]$$

$$= \frac{g_1^{n_1}\, g_2^{n_2}\, g_3^{n_3} \ldots}{n_1!\, n_2!\, n_3! \ldots}$$

$$= \prod_i^N \frac{g_i^{n_i}}{n_i!} \qquad \text{(vi)}$$

Here Π is the product sign (since probability is the product of such distributions).

When the gas is in equilibrium, the probability is maximum. When P is maximum $\log_e P$ is maximum. Hence the most probable distribution can be obtained by evaluating the maximum value of $\log_e P$ in eqn. (vi). This should also satisfy the two conditions that

$$\sum_i n_i = N \qquad \text{(vii)}$$

and $$\sum_i n_i E_i = E \qquad \text{(viii)}$$

Taking logarithm of eqn. (vi)

$$\log_e P = \sum (n_i \log_e g_i - \log_e n_i!)$$

By Stirling's theorem

$$\log_e n! = n \log_e n - n$$

$$\therefore \quad \log_e P = \sum [n_i \log_e g_i - (n_i \log_e n_i - n_i)$$

$$= \sum_i n_i - \sum (n_i \log_e n_i - n_i \log_e g_i)$$

$$\therefore \quad \log_e P = N - \sum n_i \log_e \left(\frac{n_i}{g_i} \right) \qquad \text{(ix)}$$

Differentiating eqn. (ix)

$$d(\log_e P) = -\sum_i (dn_i) \log_e \frac{n_i}{g_i} - \sum_i n_i d(\log_e \frac{n_i}{g_i})$$

$$= -\sum_i dn_i \log_e \frac{n_i}{g_i} - \sum_i n_i \frac{dn_i}{n_i}$$

$$= -\sum_i dn_i \log_e \frac{n_i}{g_i} - \sum_i dn_i$$

But $\sum_i dn_i = 0$

$$\therefore \quad -d(\log_e P) = \sum_i \left[\log_e \left(\frac{n_i}{g_i} \right) \right] dn_i \qquad \text{(x)}$$

For P to be maximum, $d(\log_e P) = 0$

$$\therefore \quad \sum_i \left[\log_e \left(\frac{n_i}{g_i} \right) \right] dn_i = 0 \qquad \text{(xi)}$$

But $\sum dn_i = 0$ (xii)

and $\sum E_i dn_i = 0$ (xiii)

Equations (xii) and (xiii) can be incorporated into eqn.(xi) by making use of *Lagrange's method of undetermined multipliers.* Multiplying eqn. (xii) by α and eqn. (xiii) by β and adding to eqn. (xi), we get

$$\sum_i \left[\log_e \left(\frac{n_i}{g_i} \right) + \alpha + \beta E_i \right] dn_i = 0$$

or $\quad \log_e \frac{n_i}{g_i} + \alpha + \beta E_i = 0$

or $\quad \frac{n_i}{g_i} = e^{-\alpha - \beta E_i}$ (xiv)

Eqn. (xiv) gives the distribution of molecules over the various energy states for the most probable configuration and α and β are two parameters that depend upon the physical property of the system.

$$N = n_1 + n_2 + n_3 + \ldots$$

$$N = g_1 e^{-\alpha-\beta E_1} + g_2 e^{-\alpha-\beta E_2} + \ldots$$

$$N = e^{-\alpha}\left[g_1 e^{-\beta E_1} + g_2 e^{-\beta E_2} + \ldots\right]$$

$$N = e^{-\alpha}\left[\sum_i g_i e^{-\beta E_i}\right]$$

Let $\sum_i g_i e^{-\beta E_i} = Z$ (xv)

where Z is called partition function. Then

$$N = e^{-\alpha}(Z) \qquad \text{(xvi)}$$

or $e^{-\alpha} = \dfrac{N}{Z}$

Substituting this value in eqn. (xiv),

$$n_i = \left(\frac{N}{Z}\right)\left[g_i\, e^{-\beta E_i}\right] \qquad (17.5)$$

Eqn. (17.5) is known as Maxwell-Boltzmann distribution law.

Maxwell-Boltzmann Distribution in terms of Temperature

The total energy of an isolated system is given by

$$E = \sum_i n_i E_i$$

$$E = n_1 E_1 + n_2 E_2 + n_3 E_3 + \ldots$$

$$E = \left(g_1 e^{-\alpha-\beta E_1}\right)E_1 + \left(g_2 e^{-\alpha-\beta E_2}\right)E_2 + \ldots$$

$$E = e^{-\alpha}\left[g_1 E_1 e^{-\beta E_1} + g_2 E_2 e^{-\beta E_2} + \ldots\right]$$

But $e^{-\alpha} = \dfrac{N}{Z}$

$$\therefore \quad E = \frac{N}{Z}\left[g_1 E_1 e^{-\beta E_1} + g_2 E_2 e^{-\beta E_2} + \ldots\right]$$

$$=\frac{N}{Z}\sum_i g_i E_i e^{-\beta E_i} \qquad \text{(i)}$$

Here $Z=\sum_i g_i E_i e^{-\beta E_i}$

$$\therefore \quad \frac{dZ}{d\beta}=\frac{d}{d\beta}\sum_i g_i e^{-\beta E_i}=-\sum_i g_i E_i e^{-\beta E_i}$$

Substituting this value in eqn. (i),

$$E=-\frac{N}{Z}\frac{d}{d\beta}\sum g_i e^{-\beta E_i}$$

$$=-\frac{N}{Z}\frac{dZ}{d\beta}$$

$$=-N\frac{d}{d\beta}[\log_e Z] \qquad \text{(ii)}$$

The average energy of a particle is given by

$$E_{av}=\frac{E}{N}=-\frac{d}{d\beta}(\log_e Z) \qquad \text{(iii)}$$

This shows that for a given system, the total energy E, the partition function Z, and the average energy of the particle E_{av}, depend on the parameter β. Therefore, β may be taken to characterize the internal energy of the system and has units per joule. If T is the temperature in degree Kelvin, then it is more customary to represent parameter β as

$$\beta=\frac{1}{kT} \quad \text{or} \quad kT=\frac{1}{\beta} \qquad \text{(iv)}$$

Here k is the Boltzmann constant and kT has the unit of energy. The unit of k is J/K and its value is given by

$$k = 1.3805 \times 10^{-23} \text{ J/K}$$

Substituting $\beta=\frac{1}{kT}$ in all equations, we get

$$Z = \sum_i g_i e^{-(E_i/kT)} \quad \text{(v)}$$

or
$$n_i = \frac{N}{Z} g_i e^{-(E_i/kT)} \quad (17.6)$$

Eqn. (17.6) represents the Maxwell-Boltzmann distribution law in terms of the temperature of the system.

$$\text{Now } \beta = \frac{1}{kT}; \text{ so } d\beta = \frac{dT}{kT^2}$$

Substituting this value in eqn. (ii)

$$E = \left(kNT^2\right)\frac{d}{dT}(\log_e Z) \quad (17.7)$$

Also, from eqn. (iii)

$$E_{av} = (kT^2)\frac{d}{dT}(\log_e Z) \quad (17.8)$$

Eqn. (17.8) gives the relation between the average energy of the particle and its temperature under equilibrium condition. Thus, the temperature of a system in statistical equilibrium can be regarded as the average energy of the particle of the system.

Maxwell-Boltzmann distribution and Molecular Energies in an Ideal Gas

The M.B distribution law is given by

$$n_i = g_i e^{-\alpha-\beta E_i} \quad \text{(i)}$$

Let us consider an ideal gas containing N molecules. If the gas has a continuous distribution of molecular energies instead of the discrete sets $E_1, E_2, E_3, \ldots$, then eqn. (i) becomes

$$n(E)dE = g(E)e^{-\alpha-\beta E}dE \quad \text{(ii)}$$

Here n(E)dE represents the number of molecules having energies between E and E+dE and g(E)dE represents the number of states having energies between E and E+dE.

Let us first find g(E)dE. A molecule of energy E has a momentum whose magnitude is given by

$$p = \sqrt{2mE} = \sqrt{p_x^2 + p_y^2 + p_z^2} \quad \text{(iii)}$$

Each set of momentum components p_x, p_y, p_z specifies a different state of motion. Let us imagine a momentum space whose coordinate axes are p_x, p_y, p_z (Fig. 17.3). The available number of

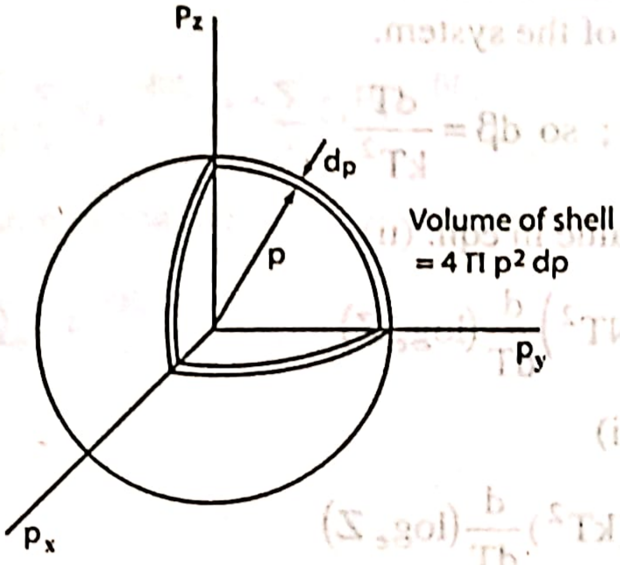


Fig. 17.3

momentum states of a particle between p and p + dp is proportional to the volume of a spherical shell in momentum space of radius p and thickness dp. Volume of this spherical shell is $4\pi p^2 dp$. Hence

$$g(p)\,dp = Bp^2dp \quad \text{(iv)}$$

where B is some constant. Since each momentum p corresponds to a single energy E, the number of energy states g(E)dE between E and E+dE is the same as the number of momentum states g(p)dp between p and p+dp.

$$\therefore \quad g(E)\,dE = Bp^2dp \quad \text{(v)}$$

But $p^2 = 2mE$ and $dp = \dfrac{mdE}{\sqrt{2mE}}$

$$g(E)dE = \sqrt{2}m^{3/2}B\sqrt{E}dE \quad \text{(vi)}$$

Substituting for g(E)dE in eqn. (ii)

$$n(E)dE = C\sqrt{E}e^{-\beta E}dE \quad \text{(vii)}$$

where $C = \sqrt{2}m^{3/2}Be^{-\alpha}$ = a constant.

To find C, use is made of normalization condition that the total number of molecules is N. Therefore,

$$N = \int_0^\infty n(E)dE = C\int_0^\infty \sqrt{E}e^{-\beta E}dE \qquad \text{(viii)}$$

From the table of definite integrals we find that

$$\int_0^\infty \sqrt{x}e^{-ax}dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

$$\therefore \quad N = C\times\frac{1}{2\beta}\sqrt{\frac{\pi}{\beta}} = \frac{C}{2}\frac{\sqrt{\pi}}{\beta^{3/2}} = C\times\frac{\sqrt{\pi}}{2}\times(kT)^{3/2}$$

$$\text{or} \quad C = \frac{2\pi N}{(\pi kT)^{3/2}} \qquad \text{(ix)}$$

$$\therefore \quad n(E)dE = \frac{2\pi N}{(\pi kT)^{3/2}}\sqrt{E}e^{-E/kT}dE \qquad \text{(x)}$$

The total internal energy of the system is

$$E = \int_0^\infty E\, n(E)dE = \frac{2\pi N}{(\pi kT)^{3/2}}\int_0^\infty E^{3/2}e^{-E/kT}dE$$

The value of the definite integral is $\frac{3}{4}(kT)^2\sqrt{\pi kT}$

$$\therefore \quad E = \frac{2\pi N}{(\pi kT)^{3/2}}\times\frac{3}{4}(kT)^2\sqrt{\pi kT} = \frac{3}{2}NkT \qquad \text{(xi)}$$

$\therefore$ The average energy of an ideal gas molecule is

$$E_{av} = \frac{E}{N} = \frac{3}{2}kT \qquad (17.9)$$

Maxwell-Boltzmann Velocity Distribution Law

The number of molecules with speeds between v and v + dv in an assembly of ideal gas containing N molecules at absolute temperature T

can be found by substituting $E = \frac{1}{2}mv^2$ and $dE = mvdv$ in eqn. (x). Then we get

$$n(v)dv = \frac{\sqrt{2}\pi\, Nm^{3/2}}{(\pi kT)^{3/2}} v^2 e^{-mv^2/2kT} dv \qquad (17.10)$$

Quantum Statistics

Many ordinarily observed phenomena such as temperature, pressure, energy, etc. could be successfully interpreted with the help of classical statistics *i.e.*, Maxwell-Boltzmann distribution law. However, it failed to explain several other experimentally ovserved phenomena such as black body radiation, photoelectric effect, specific heat capacity at low temperatures, etc. This failure of classical statistics led to the development of quantum statistics in which the discrete exchange of energy between systems instead of continuous distribution of energy was considered. There are two types of quantum statistics:

1. **Bose-Einstein Statistics:** This is applicable to identical, indistinguishable particles of zero or integral spin called *bosons.* Helium atoms at low temperature and photons are examples of bosons.
2. **Fermi-Dirac Statistics:** This is applicable to the identical, indistinguishable particles of half-integral spin obeying Pauli exclusion principle. These particles are called *fermions.* Electrons, protons, neutrons, etc, are examples of fermions. In such system of particles, no more than one particle can be in one quantum state.

The essential difference between the three statistics may be illustrated in the following simple manner. Let there be only two particles of a collection and only two cells to be occupied. The essential difference between the three statistics is illustrated below:

Maxwell-Boltzmann distribution	[p | q]	[q | p]	[pq |]	[| pq]
Bose-Einstein Statistics	[• | •]	[•• |]	[| ••]	
Fermi-Dirac Statistics	[• | •]			

17.6 Bose-Einstein Statistics

Consider an assembly of N bosons. They are identical and indistinguishable. No restriction is imposed as to how many of the particles can occupy a particular quantum state, *i.e.*, a given cell. Let us now consider that n_i indistinguishable particles are to be distributed in a box divided into g_i sections by $(g_i - 1)$ partitions. The total number of possible distributions of the particles is given by the simultaneous permutations of n_i particles and $(g_i - 1)$ partitions and is given by $(n_i + g_i - 1)!$. But this includes also the permutations of n_i particles among themselves as both these groups are internally indistinguishable. Thus the actual number of ways in which n_i particles can be distributed in g_i sublevels is

$$\frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}$$

The total number of distinguishable and distinct ways of arranging N particles in all the available energy states is given by

$$P = \frac{(n_1 + g_1 - 1)}{n_1!(g_1 - 1)!} \times \frac{(n_2 + g_2 - 1)}{n_2!(g_2 - 1)!} \times \ldots$$

$$P = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \qquad \text{(i)}$$

The most probable distribution can be obtained by finding the maximum value of $\log_e P$.

From eqn. (i),

$$\log_e P = \sum_i \left[\log_e (n_i + g_i - 1)! - \log_e n_i! - \log_e (g_i - 1)!\right]$$

According to Stirling's approximation

$$\log_e x! = x \log_e x - x$$

Also $\sum_i n_i = N$ ... ... (ii) and $\sum_i n_i E_i = E$ (iii)

$$\therefore \quad \log_e P = \sum_i [(n_i + g_i - 1)\log_e (n_i + g_i - 1) - (n_i + g_i - 1)$$

$$- (n_i \log_e n_i - n_i) - (g_i - 1)\log_e (g_i - 1) - (g_i - 1)]$$

$$= \sum\left[(n_i + g_i - 1)\log_e(n_i + g_i - 1)\right.$$

$$\left. - n_i \log_e n_i - (g_i - 1)\log_e(g_i - 1)\right] \quad \text{(iv)}$$

The maximum value of P is obtained by taking

$$d(\log_e P) = 0$$

Differentiating eqn.(iv) and setting it equal to zero

$$d(\log_e P) = \sum_i \left[\log_e(n_i + g_i - 1)dn_i - \log_e n_i dn_i\right] = 0$$

$$\therefore \quad -d(\log_e P) = \sum_i \left[-\log_e(n_i + g_i - 1)dn_i + \log_e n_i dn_i\right] = 0 \quad \text{(v)}$$

The total number of particles and the total energy are constants. So

$$\sum_i dn_i = 0 \quad \text{(vi)}$$

$$\sum_i E_i dn_i = 0 \quad \text{(vii)}$$

Multiplying eqn. (vi) by α and eqn. (vii) by β and adding to eqn. (v) we get

$$\sum_i \left[-\log_e(n_i + g_i - 1) + \log_e n_i + \alpha + \beta E_i\right]dn_i = 0$$

$$\therefore \quad -\log_e(n_i + g_i - 1) + \log_e n_i + \alpha + \beta E_i = 0$$

n_i and g_i are very large numbers compared to 1. Hence neglecting 1 we get

$$-\log_e(n_i + g_i) + \log_e n_i + \alpha + \beta E_i = 0$$

$$\log_e\left(\frac{n_i}{n_i + g_i}\right) = -\alpha - \beta E_i$$

$$\frac{n_i}{n_i + g_i} = e^{-\alpha - \beta E_i}$$

$$\text{or,} \quad \frac{n_i + g_i}{n_i} = e^{\alpha + \beta E_i}$$

$$\text{or,}\quad 1+\frac{g_i}{n_i}=e^{\alpha+\beta E_i}$$

$$\text{or,}\quad \frac{g_i}{n_i}=e^{\alpha+\beta E_i}-1$$

$$\text{or,}\quad n_i=\frac{g_i}{\left(e^{\alpha+\beta E_i}\right)-1} \qquad (17.10)$$

Eqn. (17.10) is known as Bose-Einstein distribution law.

Taking $\beta=\frac{1}{kT}$,

$$n_i=\frac{g_i}{\left(e^{\alpha+E/kT}\right)-1} \qquad (17.11)$$

Value of the constant α is governed by the eqn. $\sum n_i = N$. As n_i cannot be negative, α must always have a positive value.

Photon gas - Planck's law

Planck's formula for black body radiation may be derived from Bose-Einstein's statistics. Electromagnetic radiations trapped in a cavity and in thermal equilibrium with the walls of the cavity are termed as black body radiations. Radiations from a black body at absolute temperature T and in thermal equilibrium consists of light quanta called photons. Hence the black body radiation may be considered as a *photon gas.* It is assumed that the photons do not interact among themselves. Photons interact only with the walls of the cavity. It is further assumed that the photons are indistinguishable and many photons may have the same energy.

Each photon has an energy $h\nu$ and momentum h/ν. Photons are taken as bosons and they obey Bose-Einstein's statistics. As the photons can either be emitted or absorbed by the atoms of the wall of the cavity, the number of photons in a system is not necessarily constant. Hence the condition $\sum dn_i = 0$ is no longer valid. Due to this reason, $\alpha = 0$ in the B.E. distribution law. Further $\beta = 1/kT$. Therefore the B.E. distribution law becomes

$$n_i = \frac{g_i}{\left(e^{E/kT}\right)-1} \quad \dots\dots\dots \text{(i)}$$

In case the cavity is large as compared to the wavelength of the radiations, the energy spectrum of the photons is taken to be continuous. In this case, the energy difference between successive allowed energy states is very small. Thus replacing g_i by $g(E)dE$ and n_i by dn

$$dn = \frac{g(E)dE}{\left(e^{E/kT}\right)-1} \quad \dots\dots\dots \text{(ii)}$$

Since $E = h\nu$, $g(E)dE = g(\nu)d\nu$. The factor $g(\nu)d\nu$ corresponds to the number of oscillatory modes in the frequency range $d\nu$ and relating to the energy range dE.

The total number of states between frequencies ν and $\nu + d\nu$ can be obtained by calculating the spherical volume bound by the spheres of radii $\frac{h(\nu + d\nu)}{c}$ and $\frac{h\nu}{c}$.

The volume of the spherical shell

$$V = \frac{4}{3}\pi\frac{h^3}{c^3}(\nu + d\nu)^3 - \frac{4}{3}\pi\frac{h^3}{c^3}\nu^3$$

$$= \frac{4}{3}\pi\frac{h^3}{c^3}\left[\nu^3 + 3\nu^2 d\nu + \dots - \nu^3\right]$$

$$= \frac{4}{3}\pi\frac{h^3}{c^3}\times 3\nu^3 d\nu$$

$$= 4\pi\frac{h^3}{c^3}\nu^2 d\nu \quad \dots\dots\dots \text{(iii)}$$

The phase space has volume $V = h^3$. The number of states in the black body radiation in the frequency range ν and $\nu + d\nu$ is given by

$$g(\nu)d\nu = 4\pi\frac{V}{c^3}\nu^2 d\nu$$

Taking into account the doubling of the states due to polarization of the photons, the total number of states becomes

$$g(\nu)d\nu = 8\pi\frac{V}{c^3}\nu^2 d\nu \quad \ldots\ldots\ldots(iv)$$

$$\therefore \quad dn = \left(\frac{8\pi V\nu^2 d\nu}{c^3}\right)\times\frac{1}{\left(e^{h\nu/kT}\right)-1}$$

$$\text{or} \quad \frac{dn}{V} = \frac{8\pi\nu^2}{c^3}\cdot\frac{d\nu}{\left(e^{h\nu/kT}\right)-1} \quad \ldots\ldots\ldots(v)$$

Eqn. (v) represents the number of photons per unit volume lying in the frequency range ν and $\nu + d\nu$.

The *energy density* of radiation ($E_\nu d\nu$) of frequencies between ν and $\nu + d\nu$ is found by multiplying eqn. (v) by the energy of the photon $h\nu$.

$$\therefore \quad E_\nu d\nu = \frac{8\pi h\nu^3}{c^3}\cdot\frac{d\nu}{\left(e^{h\nu/kT}\right)-1} \qquad (17.12)$$

This is Planck's law of radiation in terms of frequency.

$$\text{But } \nu = \frac{c}{\lambda}$$

$$\text{or} \quad d\nu = -\frac{c}{\lambda^2}d\lambda$$

Neglecting the negative sign, we get

$$E(\lambda)d\lambda = \frac{8\pi h\nu^3}{c^3}\times\frac{c}{\lambda^2}\left[\frac{1}{\left(e^{h\nu/kT}\right)-1}\right]d\lambda$$

$$= \frac{8\pi hc}{\lambda^5}\left[\frac{1}{\left(e^{h\nu/kT}\right)-1}\right]d\lambda$$

$$\text{or} \quad E_\lambda = \frac{8\pi hc}{\lambda^5}\left[\frac{1}{\left(e^{h\nu/kT}\right)-1}\right] \qquad (17.13)$$

Eqn. 17.13 gives the energy density for wavelength λ in the spectrum of the black body.

Both eqns. 17.12 and 17.13 represent Planck's law for black body radiation.

17.7 Fermi-Dirac Statistics

Consider an assembly of N *fermions*. These are *identical* and *indistinguishable* particles of half-integral spin. They obey Pauli's exclusion principle. It means that no two particles can occupy the same energy level and the wave function of the whole system must be antisymmetric.

Consider N fermions with the total energy E. Suppose that n_1 particles occupy the first energy level with energy E_1, n_2 particles occupy the second energy level with energy E_2, and so on. Let us now find out the total number of ways in which n_i particles can be distributed in g_i cells having the same energy E_i. The first particle can be placed in any one of the available g_i states *i.e.*, this particle can be assigned to any one of the g_i sets of quantum number. Thus the particle can be distributed in g_i different ways. Similarly, the second particle can be arranged in $(g_i - 1)$ different ways and the process continues.

Thus the total number of different ways in which n_i particles may be distributed among the available g_i states with energy level E_i is

$$= g_i(g_i - 1)(g_i - 2)..........[g_i - (n_i - 1)]$$

$$= \frac{g_i!}{(g_i - n_i)!} \quad .... \text{(i)}$$

Further, since the particles are taken to be indistinguishable, it will not be possible to detect any difference if the n_i particles are reshuffled into different states occupied by them in the energy level E_i. Therefore, the total number of distinguishable arrangements of n_i particles in g_i states is

$$\frac{g_i!}{n_i!(g_i - n_i)!} \quad .... \text{(ii)}$$

Therefore, the total number of different and distinguishable ways in which n_1, n_2, n_3, etc. particles may be distributed among the various energy levels E_1, E_2, E_3, etc. can be obtained by

$$P = \frac{g_1!}{n_1!(g_1 - n_1)!} \cdot \frac{g_2!}{n_2!(g_2 - n_2)!} \cdots \cdots$$

$$= \prod_i \frac{g_i!}{n_i!(g_i - n_i)!} \cdots \cdots \cdots \text{(iii)}$$

The most probable distribution can be obtained by evaluating the maximum value of $\log_e P$.

This should also satisfy the condition that

$$\sum_i n_i = N \text{ and } \sum_i n_i E_i = E$$

From eqn. (iii)

$$\log_e P = \sum [\log_e g_i! - \log_e n_i! - \log_e (g_i - n_i)!]$$

Applying Stirling's approximation $[\log x! = x \log_e x - x]$

$$\log_e P = \sum [(g_i \log_e g_i - g_i) - (n_i \log_e n_i - n_i)$$
$$- \{(g_i - n_i) \log_e (g_i - n_i) - (g_i - n_i)\}] \cdots \cdots \cdots \text{(iv)}$$

Differentiating eqn. (iv)

$$-d \log_e P = \left[\sum_i \log_e n_i - \log_e (g_i - n_i) \right] dn_i$$

For P to be maximum, $d(\log_e P) = 0$

$$\therefore \quad \sum_i [\log_e n_i - \log_e (g_i - n_i)] dn_i = 0 \cdots \cdots \cdots \text{(v)}$$

$$\text{But} \quad \sum_i dn_i = 0 \cdots \cdots \cdots \text{(vi)}$$

$$\text{and} \quad \sum_i E_i dn_i = 0 \cdots \cdots \cdots \text{(vii)}$$

Multiplying (vi) by α and (vii) by β and adding to eqn. (v), we get

$$\sum_i [\log_e n_i - \log_e (g_i - n_i) + \alpha + \beta E_i] dn_i = 0$$

The equilibrium distribution is possible if

$$\log_e n_i - \log_e(g_i - n_i) + \alpha + \beta E_i = 0 \;\ldots\;\ldots\;\ldots \text{(viii)}$$

$$\therefore \quad \log_e \frac{n_i}{(g_i - n_i)} = -\alpha - \beta E_i$$

$$\text{or} \quad \frac{n_i}{g_i - n_i} = e^{-\alpha-\beta E_i}$$

$$\text{or} \quad \frac{g_i - n_i}{n_i} = e^{\alpha+\beta E_i}$$

$$\text{or} \quad \frac{g_i}{n_i} - 1 = e^{\alpha+\beta E_i}$$

$$\text{or} \quad \frac{g_i}{n_i} = \left(e^{\alpha+\beta E_i}\right) + 1$$

$$\text{or} \quad n_i = \frac{g_i}{\left(e^{\alpha+\beta E_i}\right) + 1} \qquad (17.14)$$

Eqn. (17.14) represents the *Fermi-Dirac Distribution law* or *Fermi-Dirac Statistics*

Substituting $\beta = \frac{1}{kT}$, eqn. (17.14) becomes

$$n_i = \frac{g_i}{\left[e^{(\alpha + E/kT)}\right] + 1} \qquad (17.15)$$

In most cases α is negative and is taken to be equal to $-\frac{E_F}{kT}$

$$\therefore \quad n_i = \frac{g_i}{\left[e^{(E_i - E_F)/kT}\right] + 1} \qquad (17.16)$$

The value of E_F is positive and is independent of temperature. E_F is called the *Fermi energy* and is defined as the *maximum kinetic energy that a free electron can have at the absolute zero of temperature.*

Electron Gas

Various properties of the metals such as electrical and thermal conductivities can be explained on the assumption that the electrons in the metals are free to move exactly like the particles of a gas. The behaviours of the electrons in the metal can be studied on the basis of F.D. Statistics by assuming that the electrons form an electron gas in the metal.

Consider a system of fermions in thermal equilibrium at a temperature T K. The number of electrons (fermions) having energies between *E* and E+dE is given by

$$n(E)dE = \frac{g(E)dE}{\left[e^{(E-E_F)/kT}\right]+1}$$

Here, g(E)dE is the number of states available to an electron in the energy interval E and E+dE. Now an electron has the spin $\pm\frac{1}{2}$. Therefore the total number of states for a system of free electrons enclosed in a volume V,

$$g(E)dE = 2\times\frac{4\pi p^2 dpV}{h^3}$$

Here, the factor 2 accounts for the two spin states of an electron. p is the momentum of an electron with an energy E.

Putting $\frac{1}{\left[e^{(E-E_F)/KT}\right]+1} = f(E)$,

we get

$$n(E)dE = \frac{8\pi Vp^2dp}{h^3}f(E)$$

The total number of electrons in the system is given by

$$N = \int_{E=0}^{\infty} n(E)dE = \int_{E=0}^{\infty} \frac{8\pi Vp^2dp}{h^3}f(E)$$

At the absolute zero of temperature,

$$f(E) = 1 \text{ for } E < E_F, \text{ and } f(E) = 0 \text{ for } E > E_F$$

$$\therefore \quad N = \int_{E=0}^{E_F} \frac{8\pi V p^2 dp}{h^3} = \frac{8\pi V}{h^3}\left[\frac{p^3}{3}\right]_{E=0}^{E=E_F}$$

If m is the mass of the electron, $p^2 = 2mE$

$\therefore \quad p = 0$ when $E = 0$, and $p = \sqrt{2mE_F}$ when $E = E_F$

$$\therefore \quad N = \frac{8\pi V}{h^3}\frac{\left(\sqrt{2mE_F}\right)^3}{3} = \frac{8\pi V}{h^3}\frac{(2mE_F)^{3/2}}{3}$$

$$\therefore \quad E_F = \frac{h^2}{2m}\left(\frac{3N}{8\pi V}\right)^{2/3} \tag{17.17}$$

Eqn. (17.17) gives us the value of fermi energy. The quantity $\frac{N}{V}$ is the density of free electrons.

Comparison of the three statistics

Maxwell-Boltzmann	Bose-Einstein	Fermi-Dirac
1. Particles are distinguishable	Paricles are indistinguishable	Particles are indistinguishable
2. Classical Particles can have any spin	Bose particles can have zero or integral spin	Fermi particles can have half integral spin
3. Particles do not obey Pauli's exclusion principle	Particles do not obey Pauli's exclusion principle	Particles obey Pauli's exclusion principle
4. M-B distribution law $n_i = \frac{g_i}{e^{(\alpha+\beta E_i)}}$	B-E distribution law $n_i = \frac{g_i}{e^{(\alpha+\beta E_i)} - 1}$	F-D distribution $n_i = \frac{g_i}{e^{(\alpha+\beta E_i)} + 1}$

EXERCISES

1. Distinguish between classical and quantum statistics.
2. State and prove Boltzmann's theorem connecting entropy and probability.
3. Derive Maxwell-Boltzmann distribution law.
4. Applying Maxwell-Boltzmann distribution law, show that the internal energy of an ideal mono-atomic gas depends only on the temperature.
5. Derive Bose-Einstein distribution law. How could it be used to obtain the Planck's formula for the black body radiation.
6. Apply Bose-Einstein distribution law to a photon gas.
7. Derive Fermi-Dirac distribution law. Deduce an expression for the energy of a Fermi gas at absolute zero. Point out its significance.
8. Distinguish between (i) Maxwell-Boltzmann, (ii) Bose-Einstein and (iii) Fermi-Dirac statistics. Give examples.

CHAPTER XVIII

THE ATOMIC NUCLEUS

The atomic nucleus was discovered by Rutherford in 1911. His experiments on α-particle scattering showed that almost the entire mass of the atom is confined within a small sphere of diameter of the order of $\approx 10^{-14}$ m and is located at the centre of the atom. This is called the nucleus of the atom. α-particle scattering further showed that the nucleus must have a positive charge. The nucleus is surrounded by the orbiting electrons. In this chapter we shall discuss some of the most basic characteristics of the nucleus, *viz.* its mass, size, shape and other externally observable properties. We shall also consider some deeper questions, such as the force that holds the nucleus together.

From alpha-decay studies, it was known that heavy nuclei can, to some extent, break up into smaller and identical constituents. It is therefore clear that a nucleus is built up of elementary particles. However, it was not known before 1932 exactly what these particles were. In that year Chadwick discovered the neutron and, since that discovery, it has been generally accepted that the nucleus is built up of *neutrons* and *protons.* Proton has a positive charge of the same magnitude as that of an electron but its mass is about 1837 times more than that of an electron. A neutron has almost the same mass as that of the proton but is electrically neutral. The proton and the neutron are considered to be two different charge states of the same particle which is called a *nucleon.*

In beta decay and induced reactions at high energies, other particles may emerge from the nucleus. However, it is now believed that these particles are created in the nucleus at the moment of emission and are therefore not to be considered as constituents of the nucleus.

A species of nucleus, known as a *nuclide,* is represented schematically by $_Z X^A$ where, Z, the *atomic number* indicates the number of protons, A the *mass number*, indicates the total number of protons plus neutrons and X is the chemical symbol of the species.

N = number of neutrons = A – Z.

As an example, the copper nucleus $_{29}Cu^{63}$ has Z=29 protons, A = 63 nucleons and N = 63 − 29 = 34 neutrons.

18.1 Classification of Nuclei

Depending on the number of protons or neutrons, atoms of different elements are classified as follows:

(i) **Isotopes:** Nuclei with the same atomic number Z but different mass numbers A are called *isotopes.* The nuclei $_{28}Ni^{58}$, $_{28}Ni^{60}$, $_{28}Ni^{61}$, $_{28}Ni^{62}$, $_{28}Ni^{64}$ are all isotopes of nickel. The isotopes of an element have the same number of protons but have different number of neutrons. Since the characteristic properties of an atom is ultimately determined by its nuclear charge, the isotopes of an element have identical chemical behaviour and differ physically only in mass.

(ii) **Isobars:** Nuclei having the same mass number A but different atomic number Z are called *isobars.* The nuclei $_{22}Ti^{50}$ and $_{20}Cr^{50}$ are isobars. The isobars are atoms of different elements and have different physical and chemical properties.

(iii) **Isotones:** Nuclei with the same number of neutrons N are called *isotones.* The nuclei $_6C^{14}$, $_7N^{15}$, $_8O^{16}$ are isotones since the number of neutrons in each case is 8.

(iv) **Isomers:** There are atoms, which have the same atomic number Z and same mass number A but differ from one another in their nuclear energy states and exihibit differences in their internal structure. One of these energy states may be an excited state with a relatively long-life, so that its decay time is directly observable. Such an excited state is called an *isomeric* (or metastable) state. Thus two nuclei of the same species (same Z and A) but capable of existing in different energy states, at least one of which is metastable, are called *isomers.*

(v) **Mirror nuclei:** Nuclei having the same mass number A, but with the proton and neutron number interchanged (that is, the number of protons in one is equal to the number of neutrons in the

other) are called *mirror nuclei*. Example are $_4Be^7$ (Z = 4 and N = 3) and $_3Li^7$ (Z = 3 and N = 4).

18.2 General properties of the nucleus

Some important properties of the nucleus are discussed below.

Nuclear size

Rutherford's experiments on α-scattering indicated that the mean radius of an atomic nucleus is of the order of 10^{-14} to 10^{-15}m, while that of the atom is about 10^{-10}m. Thus the nucleus is about 10,000 times smaller in radius than the atom.

The empirical formula for the nuclear radius is given by

$$R = R_0 A^{1/3}$$

where A is the mass number and $R_0 = 1.3\times10^{-15}$m. The nuclei are so small that the *fermi* (F) is taken as an appropriate unit of length where $1F = 10^{-15}$m. Thus the nuclear radius is given empirically by

$$R = R_0 A^{1/3} = 1.3\times10^{-15}\text{ m} \times A^{1/3} = 1.3\ A^{1/3}\text{ F.}$$

From this formula, the radius of the $_6C^{12}$ nucleus turns out to be 3F ($1.3 \times 12^{1/3}$F), that of $_{47}Ag^{107}$ to be 6.2F and that of $_{92}U^{238}$ to be 8.1F. More accurate measurement of nuclear radius has put the value of R_0 to be equal to

$$R = 1.2 \times 10^{-15}\text{ m} = 1.2\text{ F}$$

It should be remembered that nuclei do not have sharp boundaries. Despite this, the value of R_0 as quoted above is representative of effective nuclear sizes. From this formula the radius of $_6C^{12}$ nucleus turns out to be

$$R = R_0 A^{1/3} = 1.2\times10^{-15}\times12^{1/3}\text{ m} = 2.7\text{ F.}$$

Similarly the radius of $_{47}Ag^{107}$ nucleus is 5.7 F and that of $_{92}U^{238}$ nucleus is 7.4 F.

R. Hofstadter and his coworkers have measured the nuclear sizes more accurately by scattering high-energy electrons off various target elements throughout the periodic table. Their experiments

indicate that the *mean electromagnetic radius* of the nucleus defined as the radius of the 50 – percent point in the density distribution is given by

$$R_e = (1.07 \pm 0.02)\, A^{1/3} \times 10^{-15}\ \text{m} = 1.07\, A^{1/3}\ \text{m}.$$

Another result obtained from their experiments is that the *surface thickness,* defined as the distance between the 10-percent and 90-percent points in the density distribution is the same for all nuclei and is given by

$$t = (2.4 \pm 0.3) \times 10^{-15}\text{m}.$$

Various other methods for the nuclear radius R yield slightly varying values for the constant in front of $A^{1/3}$. However, the dependency upon A is the cube root, which reflects the fact that the density of nucleon at the centre of the nucleus is approximately the same for all nuclei.

Example 18.1 The radius of H_0^{165} *is 7.731 fermi. Deduce the radius of* He^4.

Soln.

Let R_1 and A_1 be the radius and mass number of H_0^{165} and R_2 and A_2 be the those for He^4 respectively. Then

$$R_1 = R_0 A_1^{1/3} \quad \text{and} \quad R_2 = R_0 A_2^{1/3}$$

$$\therefore \quad \frac{R_1}{R_2} = \frac{R_0 A_1^{1/3}}{R_0 A_2^{1/3}} = \frac{A_1^{1/3}}{A_2^{1/3}}$$

$$\therefore \quad R_2 = \frac{R_1 A_2^{1/3}}{A_1^{1/3}} = \frac{7.731 \times (4)^{1/3}}{(165)^{1/3}} = 2.238\text{F}$$

Nuclear mass

We know that the nucleus consists of protons and neutrons. Strictly speaking, the nuclear mass or the mass of the nucleus should be

$$Zm_p + Nm_n$$

where m_p and m_n are the masses of proton and neutron respectively while Z and N are the respective numbers of protons and neutrons in the nucleus. Nuclear mass is about 99.975% of the mass of the whole atom and is often referred to as the mass or weight of the whole atom or atomic weight. Nuclear masses are experimentally measured accurately by mass spectrometers. Nuclear masses can also be obtained from an atomic mass table by subtracting the electron masses. Except for the nuclear 'particles' which are ionised hydrogen, helium or heavier atoms, *atomic*, rather than nuclear masses, are almost always used in nuclear physics.

Nuclear density

Since the size of the nucleus is extremely small and its mass very large, its density is unbelievably high. The nuclear density ρ_N can be calculated from $\rho_N = \dfrac{\text{Nuclear mass}}{\text{Nuclear volume}}$

Now, Nuclear mass = Am_N where A = mass number and m_N = mass of a nucleon

$$\cong 1.67 \times 10^{-27} \text{ kg.}$$

$$\text{Nuclear volume} = \frac{4}{3}\pi R^3 = \frac{4}{3}\pi\left(R_0 A^{1/3}\right)^3 = \frac{4}{3}\pi R_0^3 A$$

$$\therefore \quad \rho_N = \frac{A.\left(1.67\times10^{-27}\right)}{\frac{4}{3}\pi\left(1.3\times10^{-15}\right)^3 .A} = 1.816\times10^{17} \text{ kg/m}^3$$

This is an extremely high value. This shows that the nuclear matter is in an extremely compressed state.

Nuclear charge

The charge of the nucleus is due to the protons contained in it. Each proton has a positive charge of 1.6×10^{-19}C and is the same as the charge of an electron. Therefore, the nuclear charge is Ze where Z is the atomic number of the nucleus. The value of Z is determined

from X-ray scattering experiments, from nuclear scattering of α-particles, and from the X-ray spectrum.

Nuclear quantum states

The study of α- and γ-ray spectra as well as artificial radioactivity shows that every nucleus possesses a set of quantum states and a corresponding number of discrete energy levels. Transition between different nuclear energy states are accompanied by emission of γ-rays. In this respect, there is great resemblance between quantum states inside the nucleus and those outside the nucleus *i.e.*, the quantum states of the orbital electrons. However there is one fundamental difference. In the region outside the nucleus, Coulomb's law of inverse square reigns supreme whereas inside the nucleus, short range nuclear forces come into play.

Spin angular momentum

Both the proton and neutron like the electron, have an intrinsic spin. The *spin angular momentum* is given by $L_s = \sqrt{l(l+1)}.\frac{h}{2\pi}$ where the quantum number l, commonly called the spin, is equal to $\frac{1}{2}$. Thus the spin angular momentum has a value $L_s = \frac{\sqrt{3}}{2}.\frac{h}{2\pi}$.

In addition to the spin angular momentum, the protons and neutrons in the nucleus have an orbital angular momentum. The resultant angular momentum of the nucleus is then obtained by adding the spin and orbital angular momenta of all the nucleons within the nucleus. The total angular momentum of the nucleus is given by

$$L_N = \sqrt{l_N(l_N+1)}.\frac{h}{2\pi}$$

This total angular momentum is called *nuclear spin* and is usually denoted by I.

Nuclear magnetic moment

A spinning electron has a magnetic dipole moment associated with it. This magnetic dipole moment is given by 1 Bohr magneton, *i.e.*, $\mu_e = \dfrac{eh/2\pi}{2m_e}$.

Proton has a positive charge and should have magnetic dipole moment associated with it. According to Dirac's theory, the magnetic dipole moment of a spinning proton is given by

$$\mu_N = \frac{eh/2\pi}{2m_p}$$

where m_p is the mass of the proton. μ_N is called a *nuclear magneton* and is the unit of nuclear magnetic moment. μ_N has a value of 5.050 $\times 10^{-27}$ J/T. Since m_p is = 1836 m_e, the nuclear magneton is only $\dfrac{1}{1836}$ of a Bohr magneton.

Electric quadrupole moment

In addition to magnetic moment, a nucleus may have an electric quadrupole moment. The electric dipole moment is zero for atoms and nuclei in stationary states. This is a consequence of the symmetry of nuclei about the centre of mass. However, the symmetry does not necessarily need to be spherical; there is nothing precluding the nucleus from assuring the shape of an ellipsoid of rotation, for instance. Indeed, most nuclei do assume approximately such a shape. This deviation from spherical symmetry is expressed by a quantity called the *electric quadrupole moment* and is given by

$$Q = \left(\frac{1}{e}\right)\int\left(3z^2 - r^2\right)\rho.d\tau$$

where ρ is the charge density in the nucleus. Q is actually a measure of the eccentricity of the ellipsoidal nuclear surface. Q is obviously zero for a spherically symmetric charge distribution. A charge distribution stretched in the Z-direction (prolate) will give a positive

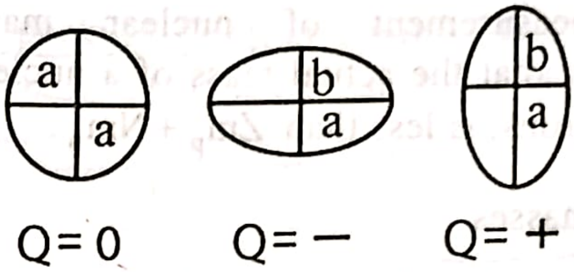


Fig. 18.1

quadrupole moment, and an oblate distribution will give a negative quadrupole moment (Fig.18.1). Since the expression for quadrupole moment is divided by the electronic charge, the dimension of the quadrupole moment is that of an area. In nuclear physics, area is measured in barns. 1 barn = 10^{-28} m^2.

18.3 Mass defect and nuclear binding energy

As already stated, the mass of a nucleus containing Z protons and N neutrons is given by $Zm_p + Nm_n$. Accurate measurement of nuclear masses by mass spectrometers, however, show that real nuclear mass $< Zm_p + Nm_n$

The unit of mass and energy in nuclear physics is the atomic mass unit (amu). Previous to 1960, the isotope of oxygen ${}_8O^{16}$ was taken as the reference for atomic mass unit. Accordingly one-sixteenth of the mass of one atom of ${}_8O^{16}$ was taken as one atomic mass unit. In 1960, the *General Assembly of the International Union of Pure and Applied Physics adopted* ${}_6C^{12}$ *as the reference for the atomic mass unit.* Now one atomic mass unit is one-twelfth of the mass of ${}_6C^{12}$ atom. The masses of protons, neutrons, etc are expressed in atomic mass unit. Carbon of atomic weight 12 and atomic number 6 has a mass equal to 12 amu.

Accordingly, 1 amu = 1.66038×10^{-27} kg

The mass of a proton = 1.007277 amu and the mass of a neutron = 1.008665 amu

Accurate measurement of nuclear masses by mass spectrometers show that the actual mass of a nucleus, containing Z protons and N neutrons, is less than $Zm_p + Nm_n$.

The difference in masses

$$Zm_p + Nm_n - \text{real nuclear mass} = \Delta m$$

is called the *mass defect*.

The mass defect of the atom is the difference of its atomic mass and its mass number.

$\therefore$ Mass defect $\Delta m = M - A$.

For carbon $_6C^{12}$ the mass defect is zero because its mass M = 12 amu and A = 12.

Although atomic masses are close to whole numbers, they almost invariably differ from integers by small amounts. This deviation of the mass of a nuclide from whole number is expressed in the form of a quantity known as *packing fraction*.

$$\text{packing fraction, } f = \frac{M - A}{A}$$

where

M = actual or isotopic mass of the nuclide in a.m.u.

A = mass number (*i.e.*, isotopic mass rounded off to the nearest integer)

The difference between atomic mass (M) and mass number (A) is the *mass defect*. Therefore,

$$\Delta m = (M - A)$$

$$\therefore \quad f = \frac{\Delta m}{A} = \frac{\text{mass defect}}{\text{mass number}}$$

Thus, packing fraction may also be defined simply as *mass defect per elementary particle in the nucleus.*

A plot of packing fraction against the corresponding mass numbers of various elements is shown in Fig.18.2. The graph shows interesting results:

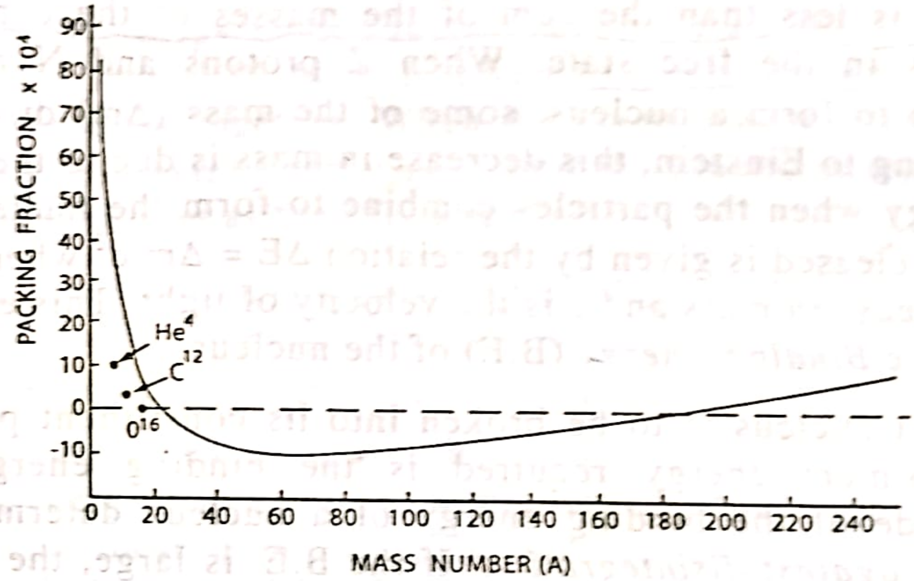


Fig. 18.2

(i) The packing fraction is positive for elements having mass number below 20.

(ii) The packing fraction is negative for elements having mass number between 20 and 200.

(iii) The packing fraction beyond mass number 200 becomes positive again and increases with increase in mass number.

The mass defect and packing fraction only show their relations with respect to carbon. The packing fraction is zero for carbon A = 12 and Z = 6.

It has been found that packing fraction is a fundamental property of a nucleus and is directly related to the availability of nuclear energy and stability. A negative packing fraction and hence mass defect means that the atomic mass of a nucleus is less than the nearest whole number. This suggests that some mass has been converted into energy during the formation of that particular nucleus. Since same amount of energy would have to be supplied in order to break up the nucleus, *a negative packing fraction implies exceptional nuclear stability*. On the other hand, a positive packing fraction (and hence mass excess) indicates that the nucleus is somewhat unstable.

19.4 Nuclear mass and binding energy

Measurement of nuclear mass shows that the mass of a nucleus is less than the sum of the masses of the constituent particles in the free state. When Z protons and N neutrons combine to form a nucleus, some of the mass (Δm) disappears. According to Einstein, this decrease in mass is due to the release of energy when the particles combine to form the nucleus. The energy released is given by the relation $\Delta E = \Delta m.c^2$ where Δm is the decrease in mass and c is the velocity of light. This energy is called the *Binding Energy* (B.E) of the nucleus.

If a nucleus is to be broken into its constituent particles, the minimum energy required is the bindiing energy. The magnitude of the binding energy of a nucleus determines its *stability against disintegration.* If the B.E. is large, the nucleus is stable.

If M is the experimentally determined mass of a nuclide having Z protons and N neutrons,

$$\text{B.E.} = \left[\left(Zm_p + Nm_n\right) - M\right]c^2$$

Let us calculate the mass defect and hence the binding energy of a deuteron which consists of one proton and one neutron.

Mass of a proton	= 1.007825 a m u
+ Mass of a neutron	= 1.008665 a m u
Expected mass of a deuteron nucleus	= 2.016490 a m u

The measured mass of a deuteron is 2.014103 amu which is 0.002387 amu *less* than the combined masses of a proton and a neutron.

It is reasonable to suppose that the *missing mass* of 0.002387 u corresponds to the energy given off when a deuteron nucleus is formed from a free proton and neutron. Since the energy equivalent of 1u of mass is 931 MeV, the energy that corresponds to the missing deuteron mass is

$$(0.002387 \text{ u})\left(931 \frac{\text{MeV}}{\text{u}}\right) = 2.22 \text{ MeV}.$$

To test this interpretation of missing mass, experiments can be performed to see how much energy is needed to break apart a deuterium nucleus into a separate neutron and proton. The required energy indeed turns out to be 2.22 MeV. When an energy less than 2.22 MeV is given to the deuterium nucleus, the nucleus stays together. When the given energy is more than 2.22 MeV, the extra energy goes into the kinetic energy of the neutron and proton as they fly apart.

Nuclear binding energies are strikingly high. The range for stable nuclei is from 2.22 MeV for deuterium to 1640 MeV for $_{83}Bi^{209}$ (an isotope of the metal bismuth). To appreciate how high the binding energies are, these may be compared with more familiar energies in terms of kilojoules of energy per kilogram of mass. In these units, a typical binding energy is 8×10^{11} kJ/kg – 800 billion kJ/kg. By contrast, to boil water involves a heat of vapourization of a mere 2260 kJ/kg, and even the heat given off by burning gasoline is only 47×10^4 kJ/kg, 17 million times smaller.

The *binding energy per nucleon* for a given nucleus is found by dividing its total binding energy by the number of nucleons (protons and neutrons) it contains. Accordingly, the binding energy per nucleon for deuterium nucleus is 2.22 MeV/2 = 1.11 MeV and for bismuth it is 1640 MeV/209 = 7.8 MeV/nucleon.

The binding energy per nucleon is plotted as a function of number of nucleons in various nuclei in Fig. 18.3. The curve rises steeply at first and then more gradually until it reaches a maximum of 8.79 MeV at A = 56 corresponding to the iron nucleus $_{26}Fe^{56}$. The curve then drops slowly to about 7.6 MeV at the highest mass numbers. It is obvious from the graph that nuclei of intermediate mass are the most stable, the most stable of them all being $_{26}Fe^{56}$, since the greatest amount of energy must be supplied to liberate each of their nucleons.

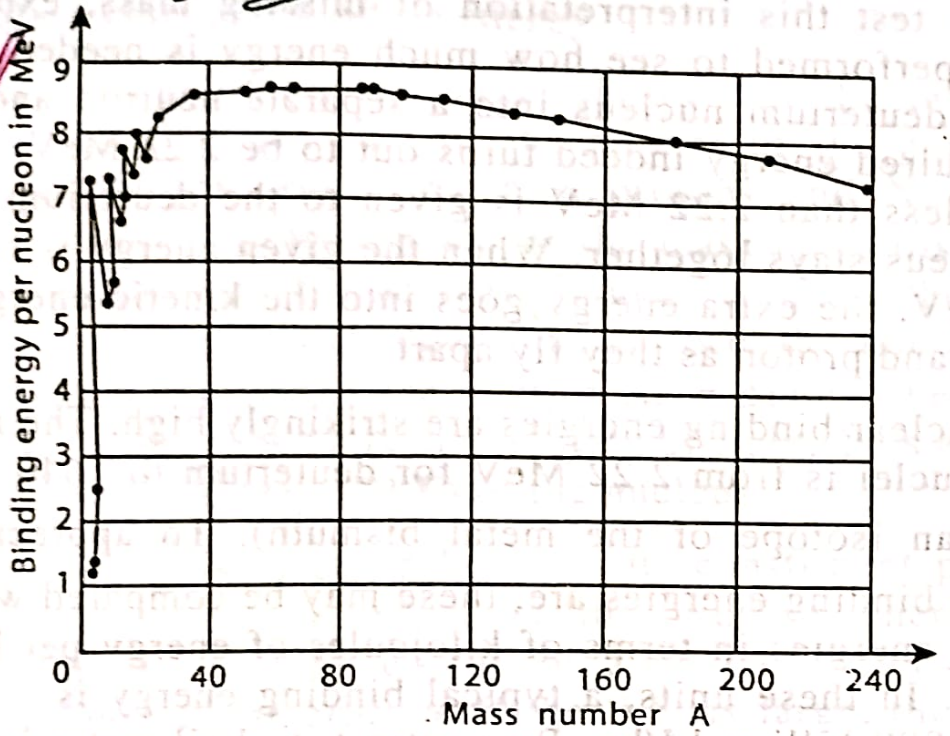


Fig. 18.3

Two remarkable conclusions can be drawn from the curve of Fig. 18.3. The first is that if a heavy nucleus is somehow split into two medium-sized nuclei, this will result in the release of a large amount of energy. For instance, if the uranium nucleus $_{92}U^{235}$ is broken into two smaller nuclei, the binding energy difference per nucleon is about 0.8 MeV. The total energy given off is therefore

$$\left(0.8\frac{\text{MeV}}{\text{nucleon}}\right)(235 \text{ nucleus}) = 188 \text{ MeV}$$

This is truly an enormous amount of energy to be produced in a single atomic event. Ordinary chemical reactions involve rearrangements of the electrons in atoms and liberate only a few electronvolts per reacting atom. Splitting a heavy nucleus into two medium sized ones, a process called *nuclear fission,* thus involves a hundred million times more energy per atom, say, the burning of coal or oil.

The other notable conclusion is that if two light nuclei can somehow be joined to form heavier ones, a large amount of energy will be liberated. For instance, if two $_1H^2$ deuterium nuclei combine

to form $_2He^4$ helium nucleus, over 23 MeV is released. Such a process, called *nuclear fusion,* is also a very effective way to obtain energy. In fact, nuclear fusion is the main energy source of the sun and the other stars.

***Example 18.2** Calculate how many times nuclear matter is denser than water if nuclear radius is given by $1.2\times10^{-15}A^{1/3}$ where A is the mass number. Assume that protons and neutrons posses equal massess.*

Soln.

Volume of a nucleus of mass number $A=\frac{4}{3}\pi R^3$

$$R= 1.2\times10^{-15}A^{1/3}$$

$$=\frac{4}{3}\pi\left(1.2\times10^{-15}A^{1/3}\right)^3$$

$$=\frac{4}{3}\pi 1.2^3\,10^{-45}\,A\ m^3$$

Let mass of a proton = mass of a neutron

$= 1.008$ a.m.u

$= 1.008\times1.66\times10^{-27}$ kg

$= 1.67\times10^{-27}$ kg.

Since the combined number of protons and neutrons, is A,

mass of nucleus $= A\times1.67\times10^{-27}$ kg

$\therefore$ Density of nucleus $=\frac{\text{mass}}{\text{volume}}$

$$=\frac{A\times1.67\times10^{-27}\,kg}{\frac{4}{3}\pi 1.2^3\times10^{-45}\,A}$$

$$= 2.3\times10^{17}\ kg/m^3$$

Density of water = 1 gm/cm^3 = 10^3 kg/m^3

$$\therefore \quad \frac{\text{density of nucleus}}{\text{density of water}} = \frac{2.3\times10^{17}\,\text{kg/m}^3}{10^3\,\text{kg/m}^3} = 2.3\times10^{14}.$$

Example 18.3 Compute the equivalent energy per gram in kwh.

Soln.

According to Einstein's mass energy relation

$E = mc^2$ $\qquad\qquad$ $m = 1\text{gm} = 10^{-3}$ kg

$c = 3 \times 10^8$ m/s

$$\therefore \quad E = \left(10^{-3}\right)\left(3\times10^8\right)^2$$

$= 9 \times 10^{13}$ Joules.

$1\ \text{kWh} = 36 \times 10^5$ J.

$\therefore$ equivalent energy per gram

$$= \frac{9\times10^{13}}{36\times10^5} = 2.5\times10^7\ \text{kWh}.$$

Example 18.4 Calculate the binding energy when

(a) *one neutron and one proton combine to form a deuteron.*

(b) *two neutrons and two protons combine to form an α-particle*

Given, mass of neutron = 1.00893 amu; mass of proton = 1.00813 amu; mass of deuteron = 2.01473 amu; mass of α-particle = 4.00389 amu.

Soln.

(a) Total mass of separate constituents of deuteron

= 1.00893 + 1.00813 = 2.01706 amu

Mass of deuteron = 2.01473 amu

$\therefore$ difference in mass Δm = (2.01706 – 2.01473) amu

= 0.00233 amu

Now 1 amu = 931 MeV

∴ Binding energy of deuteron = 931 × 0.00233 = 2.17 MeV

(b) Total mass of two protons = 2 × 1.00813 = 2.01626 amu

Total mass of two neutrons = 2 × 1.00893 = 2.01786 amu

Total mass of separate constituents of α-particle

= 2.01626 + 2.01786 = 4.03412 amu

Mass of α-particle = 4.00389 amu

Difference in mass = 4.03412 – 4.00389 = 0.03023 amu.

∴ Binding energy of α-particle

= 0.03023 × 931 = 28.1 MeV.

Example 18.5 *$_8O^{16}$ and $_8O^{18}$, the two isotopes of oxygen have nuclear masses 15.990523 amu and 17.994768 amu respectively. Calculate the binding energy per nucleon for the two isotopes. Which one of the two isotopes you expect to be more abundant given: mass of a proton = 1.007276 amu; mass of a neutron = 1.008665 amu.*

Soln.

For $_8O^{16}$ isotopes:

number of protons = 8

number of neutrons = 8

Difference in mass = 8 [$m_p + m_n$] – 15.990523

= 8[1.007276 + 1.008665] – 15.990523

= 16.127528 – 15.990523

= 0.137005 amu

∴ binding energy = 0.137005 × 931 MeV

= 127.55 MeV

$_8O^{16}$ has 16 nucleons; so binding energy per nucleon = $\frac{127.55}{16} = 7.972$ MeV.

For $_8O^{18}$ isotope:

Number of protons = 8 and number of neutrons = 10

Difference in mass = $[8\ m_p + 10\ m_n] - 17.994768$

$= [8 \times 1.007276 + 10 \times 1.008665] - 17.994768$

$= 0.15009$ amu

binding energy = 0.15009×931 MeV = 139.73 MeV.

There are 18 nucleons; so binding energy per nucleon = $\frac{139.73}{18} = 7.762$ MeV.

Since binding energy per nucleon for $_8O^{16}$ is more, $_8O^{16}$ is more stable then $_8O^{18}$.

Example 18.6 *Calculate the binding energy and packing fraction for helium. The atomic masses of proton, neutron and helium are 1.00814 amu, 1.00898 amu and 4.00387 amu and 1 amu = 931 MeV.*

Soln.

Number of protons in helium = 2

Number of neutrons in helium = 2

Difference in mass = $2\ [m_p + m_n] - m_{He}$

$= 2\ [1.00814 + 1.00898] - 4.00387$

$= 0.03037$ amu.

$\therefore$ binding energy = 0.03037×931 MeV

$= 28.275$ MeV

Mass defect, $\Delta m = M - A$ where M = mass = 4.00387

A = mass number = 4

$$\text{packing fraction} = \frac{\Delta m}{A} = \frac{M-A}{A}$$

$$= \frac{4.00387-4}{4} = 0.0009675$$

Significance of binding energy per nucleon

A remarkable fact about nuclei is that the average binding energy of a nucleon (proton or neutron) is about 8 MeV. It shows that a given nucleon does not interact with all other nucleons in the nucleus. Because, if it did, the binding energy per nucleons would have increased linearly with the number of nucleons in a nucleus on the basis of attractive forces between each pair. It is, therefore, apparent that nucleons interact only with immediate neighbours just as gas or liquid molecules interact only with their neighbouring molecules. This fact is known as *saturation* of nuclear forces and is consistent with the short range of nuclear forces.

18.5 Nuclear Forces

The nucleus consists of proton and neutrons. The question arises, as to how these protons and neutrons are kept together in the nucleus and why the mutual repulsion of the protons in the nucleus, does not disintegrate the nucleus. Hence, if stable complex nuclei are to exist, there must be some sort of attractive forces in the nuclei strong enough to overcome the repulsive forces. These attractive forces are the specific *nuclear forces* and they have the following properties.

(i) Nuclear forces are fundamentally different from the gravitational and electrostatic forces.

(ii) Nuclear forces are effective only at short ranges. Nuclear forces are appreciable only when the distance between nucleons is of the order of 10^{-15} m or less. The force vanishes for all practical purposes at distance greater than a few times 10^{-15} m. These distances are referred to as the *range* of nuclear forces. However at very short range, nuclear force turns repulsive, since otherwise the nucleons in a nucleus would mesh together.

(iii) Nuclear forces are charge independent *i.e.*, nuclear forces are the same for proton-proton, proton-neutron and neutron-neutron interactions.

(iv) Nuclear forces are spin dependent.

(v) Nuclear forces are the strongest known forces in nature. The energy needed to separate a nucleon from the nucleus is about 8 MeV in contrast to a few electron-volts required to separate the extra-nuclear electrons from the atom.

(vi) Nuclear forces have saturation property. Nuclear forces are limited in range. As a result, each nucleon interacts with only a limited number of nucleons nearest to it. This is referred to as *saturation property* of nuclear forces.

Japanese physicist Hideki Yukawa proposed in 1935 that particles intermediate in mass between electrons and nucleons are responsible for nuclear forces. Today these particles are called *pions,* a contraction of the original name π meson. Pions may be charged (π^+, π^-) or neutral (π^0), and are members of a class of elementary particles collectively called *mesons*.

According to Yukawa's theory, every nucleon continually emits and reabsorbs pions. If another nucleon is nearby, an emitted pion may shift across to it instead of returning to its parent nucleon. The associated transfer of momentum is equivalent to the action of a force.

The forces that act, between one neutron and another, and between one proton and another, are the result of the exchange of neutral mesons (π^0) between them. The force between a proton and a neutron is the result of the exchange of charged mesons (π^- and π^+) between them.

Thus a neutron emits a π^- meson and is converted into a proton:

$$n \rightarrow p + \pi^-$$

The absorption of the π^- meson by the proton (with which the neutron was interacting) converts it into a neutron:

$$p + \pi^- \rightarrow n$$

In the reverse process, a proton emits a π^+ meson, becoming a neutron and the neutron, on receiving this π^+ meson, becomes a proton:

$$p \rightarrow n + \pi^+$$

$$n + \pi^+ \rightarrow p$$

Thus in the nucleus of an atom, attractive forces exist between (i) proton and proton, (ii) proton and neutron and (iii) neutron and neutron. The forces of attraction are much stronger than the electrostatic forces of repulsion between the protons, thereby making the nucleus stable.

18.6 Nuclear Models

The short range attractive force that binds the nucleons so securely into nuclei is by far the strongest type of force known. Unfortunately the forces acting in the nucleus is nowhere near as well understood as the electromagnetic force. The nuclear structure remains unclear when compared to the atomic structure. However, even without a complete understanding of the nuclear force, considerable progress has been made in devising nuclear models able to account for prominent aspects of nuclear properties and behaviour. These models have been based on (i) the extrinsic analogy between the properties of atomic nuclei and those of a liquid drop (ii) the electron shell of an atom, etc. The corresponding models are called the liquid drop model, shell model, etc.

18.7 The Liquid Drop Model

Liquid drop model was proposed by Neils Bohr who observed that there are certain marked similarities between an atomic nucleus and a liquid drop. These similarities are as follows:

(i) In the stable state, a nucleus is supposed to be spherical in shape, just as a liquid drop is spherical due to the symmetrical surface tension forces.

(ii) Just as the force of surface tension acts on the surface of the liquid drop, there is a potential barrier at the surface of the nucleus.

(iii) The density of a liquid drop is independent of its volume. Similarly, the density of the nucleus is independent of its volume.

(iv) The intermolecular forces in a liquid are short range forces. The molecules in a liquid drop interact only with their immediate neighbours. Similarly, the nucleons in the nucleus also interact only with their immediate neighbours. That means, the nuclear forces are also short range forces. This leads to the saturation of nuclear forces and a constant binding energy per nucleon.

(v) When the temperature of the liquid is raised, the molecules from a liquid-drop evaporate due to their increased energy of thermal agitation. Similarly, when a nucleus is bombarded with nuclear projectiles, it gains energy and a compound nucleus is formed. But the compound nucleus emits nuclear radiations almost immediately.

(vi) When a small drop of liquid is allowed to oscillate, it breaks up into two smaller drops of equal size. Similarly, in nuclear fission the nucleus breaks up into two nuclei of almost equal mass numbers.

18.8 Semi-empirical mass formula

The liquid drop model can be used to obtain an expression for the binding energy of the nucleus. We shall start by assuming that the energy associated with each nucleon-nucleon bond has some value U. This energy is actually negative since attractive forces are involved, but is usually written as positive since binding energy is considered a positive quantity for convenience.

Volume energy: Since each bond energy is shared by two nucleons, the binding energy associated with each nucleon is $\frac{1}{2}U$. When an assembly of spheres of the same size is packed together in the smallest possible volume, as we suppose to be the case of nucleons inside a nucleus, each interior sphere has 12 other spheres in contact with it. Hence each interior nucleon in a nucleus has a binding energy of $(12)(\frac{1}{2}U)$ or 6U. If all the A nucleons in a nucleus could be considered to be in the interior of the nucleus, the total binding energy of the nucleus would be

$$E_v = (6U)(A) = 6\ AU\ldots$$

E_v is often written simply as

$$E_v = a_1 A \qquad \textbf{Volume energy}$$

The energy E_v is called the *volume energy* of a nucleus and is directly proportional to A.

Surface energy: Actually not all the nucleons are within the interior of the nucleus. Some nucleons are on the surface of the nucleus and therefore have fewer than 12 neighbours. The number of such nucleons depends on the surface area of the nucleus in question. A nucleus of radius R has an area of

$$4\pi R^2 = 4\pi R_0^2 A^{2/3}$$

Hence the number of nucleons with fewer than the maximum number of bonds is proportional to $A^{2/3}$. Thus the total binding energy is reduced by

$$E_s = -\, a_2 A^{2/3} \qquad \textbf{Surface energy}$$

The negative energy E_s is called the *surface energy* of a nucleus. It is most significant for the lighter nuclei since a greater fraction of their nucleons are on the surface. Because natural system always tend to evolve toward configurations of minimum potential energy, nuclei tend toward configurations of maximum binding energy. Hence a nucleus should exhibit the same surface-tension effects as a liquid drop. Therefore, in the absence of external forces it should be spherical, since a sphere has the least surface area for a given volume.

Coulomb energy

The electrostatic force of repulsion between two proton (a pair of protons) in a nucleus also decreases the binding energy of the nucleus. The *Coulomb energy* E_c of a nucleus is the work that must be done to bring together Z protons from infinity and assemble them into a spherical aggregate equal to the size of the nucleus. The potential energy of a pair of protons separated by a distance r is given by

$$V = -\frac{e^2}{4\pi\varepsilon_0 r}$$

Since there are Z(Z –1)/2 pairs of protons in a nucleus,

$$E_c = \frac{Z(Z-1)}{2}.V = -\frac{Z(Z-1)e^2}{8\pi\varepsilon_0}\left(\frac{1}{r}\right)_{av}$$

where $\left(\frac{1}{r}\right)_{av}$ is the value of $\left(\frac{1}{r}\right)$ averaged over all proton pairs. If the protons are uniformly distributed throughout a nucleus of radius R, $\left(\frac{1}{r}\right)_{av}$ is proportional to $1/R$ and hence to $1/A^{1/3}$, so that

$$E_c = -a_3\frac{Z(Z-1)}{A^{1/3}} \qquad \textbf{Coulomb energy}$$

The Coulomb energy is negative because it arises from an effect that opposes nuclear stability.

The total binding energy of the nucleus according to liquid drop model ought to be the sum of its volume, surface and coulomb energies:

$$\therefore \quad E_b = E_v + E_s + E_c = a_1A - a_2A^{2/3} - a_3\frac{Z(Z-1)}{A^{1/3}}$$

The *binding energy per nucleon* is therefore

$$\frac{E_b}{A} = a_1 - \frac{a_2}{A^{1/3}} - a_3\frac{Z(Z-1)}{A^{4/3}} \qquad (18.1)$$

Each of the terms of eqn. 18.1 is plotted in Fig. 18.4 versus A, together with their sum E_b/A. The coefficients were chosen to make the E_b/A curve resemble as closely as possible the empirical binding energy per nucleon curve (Fig. 18.3)

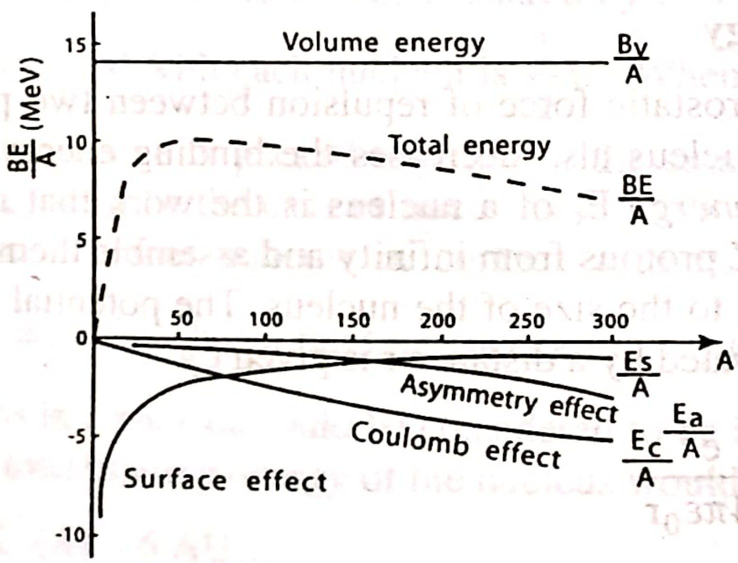


Fig. 18.4

The fact that the theoretical curve can be made to fit so well with the empirical one means that the analogy between a nucleus and a liquid drop has some validity at least.

Assymetry energy:

The binding energy formula as given by eqn. 18.1 can be improved by two effects that do not fit the simple liquid drop model but which can be readily explained in terms of a model that take into account nuclear energy levels (shell model). One of these effects occur when the neutrons in a nucleus outnumber the protons or vice versa, which means that higher energy levels have to be occupied than would be the case if N and Z are equal.

Let the uppermost neutron and proton energy levels, which the *exclusion principle* limits to two particles each, have the same spacing ε, as in Fig. 18.5. In order to produce a neutron excess of, say, $N - Z = 8$ without changing A, $\frac{1}{2}(N - Z) = 4$ protons would have to replace neutrons in an original nucleus in which $N = Z$. The new neutrons would occupy levels higher in energy by $2\varepsilon = 4\varepsilon/2$ than those of the protons they replace. In the general case of $\frac{1}{2}(N - Z)$ protons to be shifted, each

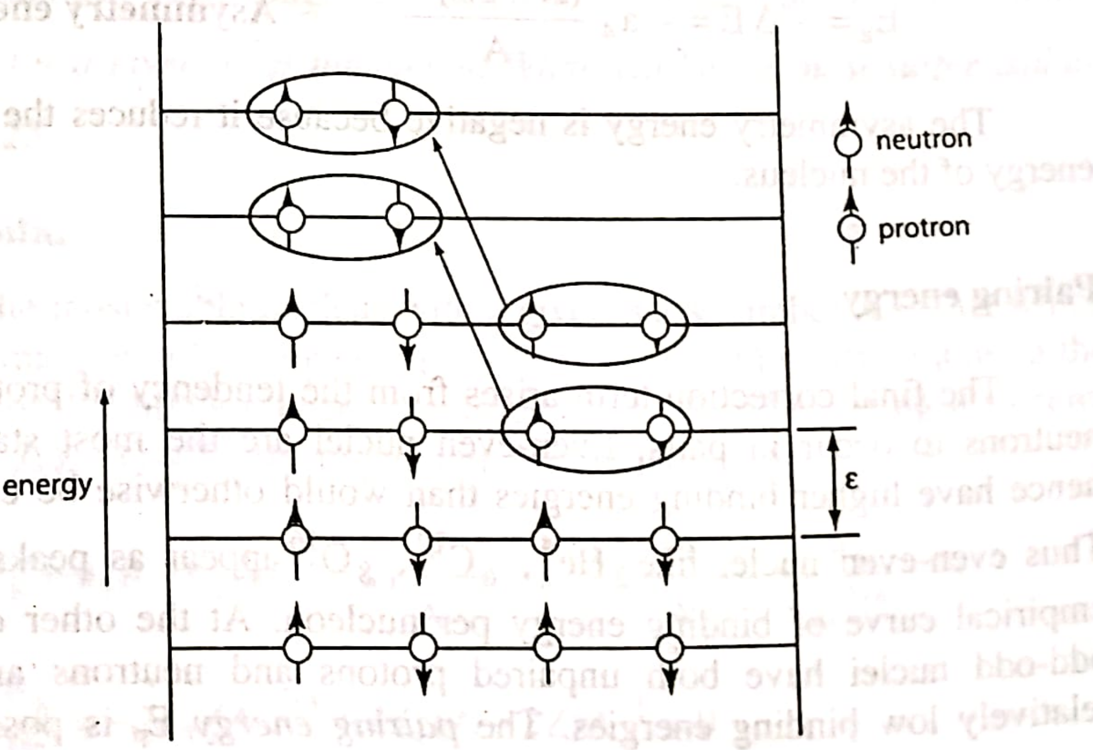


Fig. 18.5

must be raised in energy by $\frac{1}{2}(N - Z)\,\varepsilon/2$. The total work that must be done is

$$\Delta E = (\text{number of new neutrons}) \left(\frac{\text{energy increase}}{\text{new neutron}}\right)$$

$$= \left[\frac{1}{2}(N-Z)\right]\left[\frac{1}{2}(N-Z)\frac{\varepsilon}{2}\right] = \frac{\varepsilon}{8}(N-Z)^2$$

The same formula will hold if $Z > N$, since $(N-Z)^2$ is always positive. Because $N = A - Z$, $(N-Z)^2 = (A-2Z)^2$, and

$$\Delta E = \frac{\varepsilon}{8}(A-2Z)^2$$

The greater the number of nucleons in a nucleus, the smaller is the energy level spacing ε, with ε proportional to $1/A$. The departure from $N = Z$ introduces an *asymmetry* $N - Z$, which results in a decrease in binding energy. This means that the *asymmetry energy* E_a due to the difference between N and Z can be expressed as

$$E_a = -\Delta E = -a_4 \frac{(A-2Z)^2}{A} \qquad \textbf{Asymmetry energy}$$

The asymmetry energy is negative because it reduces the binding energy of the nucleus.

Pairing energy

The final correction term arises from the tendency of protons and neutrons to occur in pairs. Even-even nuclei are the most stable and hence have higher binding energies than would otherwise be expected. Thus even-even nuclei like $_2He^4$, $_6C^{12}$, $_8O^{16}$ appear as peaks on the empirical curve of binding energy per nucleon. At the other extreme, odd-odd nuclei have both unpaired protons and neutrons and have relatively low binding energies. The *pairing energy* E_p is positive for

even-even nuclei, 0 for odd-even and even-odd nuclei, and negative for odd-odd nuclei and seems to vary with A as $A^{-3/4}$. Hence

$$E_p = (\pm, 0)\frac{a_5}{A^{3/4}} \qquad \textbf{Pairing energy}$$

The final expression for the binding energy of a nucleus of atomic number Z and mass number A was first obtained by C, F. von Weizsäcker in 1935 and is known as Weizsäcker's semi-empirical mass formula. This is given by

$$E_b = a_1A - a_2A^{2/3} - a_3\frac{Z(Z-1)}{A^{1/3}} - a_2\frac{(A-2Z)^2}{A}(\pm,0)\frac{a_5}{A^{3/4}}$$

The best values of the constants, expressed in MeV, are

$$a_1 = 15.76,\ a_2 = 17.81,\ a_3 = 0.711,\ a_4 = 23.702 \text{ and } a_5 = 34$$

The contributions of the various effects in Weizsäeker's empirical formula are represented schematically in Fig. 18.4.

Example 18.7 Calculate the atomic number of the most stable nucleus for a given mass number A. What will be the most stable isobar for A = 25.

Soln.

The most stable nucleus with a given mass number A is that which has the maximum binding energy. This is obtained by differentiating the expression for binding energy with respect to Z and equating it to zero. Writing $Z(Z-1) = Z^2$, the expression for binding energy is

$$E_b = a_1A - a_2A^{2/3} - a_3\frac{Z^2}{A^{1/3}} - a_2\frac{(A-2Z)^2}{A} \pm \frac{a_5}{A^{3/4}}$$

$$\frac{dE_b}{dZ} = -2a_3ZA^{-1/3} + 4a_4(A-2Z)A^{-1} = 0$$

Introducing the numerical values a_3 and a_4, we have

$$Z \cong \frac{A}{2+0.0157A^{2/3}}$$

For light nuclei (small A), the second term in the denominator may be neglected. Then, we have $Z \cong \frac{A}{2}$.

For A =25 we have Z = 11.7 or Z ≅ 12. This nuclide is $_{12}Mg^{25}$ which is the only stable isobar with A = 25. The other isobar $_{11}Na^{25}$ and $_{13}Al^{25}$ are both radioactive.

18.9 The Shell Model

The shell model assumes that the energy structure of the nucleus (energy levels of the nucleons) is similar to that of electrons in an atom. According to this model, the protons and neutrons are grouped in shells in the nucleus in a manner similar to grouping of extranuclear electrons in various shells outside the nucleus. The shells are regarded as "filled" when they contain a specific number of protons or neutrons or both. It should be remembered that the case is not so simple as it appears. In the extranuclear shells only one type of particles (*i.e.,* electrons) is to be arranged in different shells and *Pauli's exclusion principle* is applied. In the case of the nucleus there are two types of particles, protons and neutrons and the shell arrangement is only empirical and based upon the study of the stability and interactions of the nuclides which are known.

The shell model is sometimes referred to as the *independent particle model* because it assumes that each nucleon moves independently of all the other nucleons and is acted on by an average nuclear field produced by the action of all the other nucleons.

Nuclei that have 2, 8, 20, 28, 50, 82 and 126 neutrons or protons are more abundant than other nuclei of similar mass number, suggesting that their structures are more stable. Other evidence also points up the significance in nuclear structure of the numbers 2, 8, 20, 28, 50, 82 and 126. These numbers have become known as *magic numbers*. Nuclei containing magic number nucleons of the same kind form some sort of closed nuclear shell structures. The main points in favour of this inference are:

(i) The inert gases with-closed electron shells exhibit a high degree of chemical stability. Similarly nuclides whose nuclei contain a magic number of nucleons of the same kind exhibit more than average stability.

(ii) Nuclides of isotopes of elements having an isotopic abundance greater than 60% contain magic numbers.

(iii) The stable end product, of the three main radioactive series (*viz*, the uranium, actinium and thorium series) is $_{82}Pb^{208}$ which contains 82 protons and 126 neutrons. Thus $_{82}Pb^{208}$ is the most stable isotope. This shows that the number 82 and 126 indicate stability.

(iv) Tin $(_{50}Sn)$ has *ten* stable isotopes, while $(_{20}Ca)$ has *six* stable isotopes. So elements with Z =50, 20 are usually more stable.

(v) $_{8}O^{17}$, $_{36}Kr^{87}$ and $_{54}Xe^{137}$ isotopes spontaneously emit neutrons when excited above the nucleon binding energy by a preceding β-decay. For these isotopes the number of neutrons are 9, 51 and 83 respectively which can be written as 8+1, 50+1 and 82+1. If this loosely bound neutron is interpreted as a valency neutron, then the neutron numbers 8, 50 and 82 represent greater stability than other neutron numbers.

It is apparent from the above conclusions that nuclear behaviour is often determined by the excess or deficiency of nucleons with respect to closed shells of nucleons corresponding to the magic numbers. It was, therefore, suggested that the protons and neutrons are grouped in shells inside the nucleus similar to grouping of electrons in various orbits outside the nucleus. The shells are regarded 'filled' when they contain a specific number of protons or neutrons or both. Each nucleon shell has a specific maximum capacity. When the shells are filled to capacity, they give rise to particular numbers (the magic numbers) characteristic of unusual stability.

18.10 Nuclear Fission and Fusion

It was seen in Art. 18.4 that a great deal of binding energy will be released if a large nucleus could be broken into smaller ones. But nuclei are ordinarily not at all easy to split. What we need is a way to disrupt a heavy nucleus without using more energy than that we get back from the process.

Enrico Fermi, in1934, observed that if uranium is bombarded by neutrons, several β-ray activities with different half-lives are produced. Since uranium decays by α-particle emission with a very long half-life, it was assumed that transuranic elements (Z>92) were being formed. However, Hahn and Strassman, in 1938, showed by very careful analysis that one of the radioactive elements produced when uranium is bombarded by neutrons is an isotope of barium $\left({}_{56}Ba^{141}\right)$. They found that barium was accompanied by a radioactive isotope of krypton (Z = 36). The atomic numbers of these two nuclides add up to 92 (56 +36), the atomic number of uranium. It was suggested by Meitner and Frisch that on neutron bombardment, the uranium nucleus splits up into two lighter nuclei (Ba and Kr). The phenomenon was given the name *fission*, since it resembled the division of cell in biology.

The process of breaking up of the nucleus of a heavy atom into two, more or less equal nuclei with the release of an enormous amount of energy is known as fission. The new nuclei that result from fission are called *fission fragments.*

The process actually involves two separate stages. When uranium is bombarded with neutrons, first the uranium nucleus captures a slow neutron to form a new nucleus, called a *compound nucleus,* whose atomic and mass numbers are respectively the sum of the atomic numbers of the original particles and the sum of their mass numbers.

Compound nuclei have lifetime of 10^{-16}s or so. Although too short to permit actually observing such nuclei events directly, such lifetimes are nevertheless long relative to the 10^{-21}s or so required for a nuclear particle with an energy of several MeV to pas through a nucleus.

The compound nucleus has no 'memory' of how it was formed, since its nucleus are mixed together regardless of origin and the energy brought into it by the incident particle is shared among all of them. The compound nucleus then splits into two nearly equal parts. Because heavy

nuclei have a greater neutron/proton ratio than lighter ones, the fragments contain an excess of neutron. To reduce this excess, two or three neutrons are emitted by the fragments as soon as they are formed. The schematic equation for the fission process is

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow X + Y + \text{neutrons}$$

$_{92}U^{236*}$ is a highly unstable isotope, and X and Y are the fission fragments. The fragments are not uniquely determined because various combinations of fragments along with emission of various number of neutrons are possible. Typical fission reactions are

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow {}_{56}Ba^{141} + {}_{36}Kr^{92} + 3{}_0n^1 + Q \quad (18.3)$$

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow {}_{54}Xe^{140} + {}_{38}Sr^{94} + 2{}_0n^1 + Q \quad (18.4)$$

where Q is the energy released in the reaction.

According to eqn. 19.3, when $_{92}U^{235}$ is bombarded by a slow neutron, the nucleus becomes unstable and begins to oscillate violently, splitting into $_{56}Ba^{141}$ and $_{36}Kr^{92}$ with the emission of three neutrons and energy Q (Fig. 18.6).

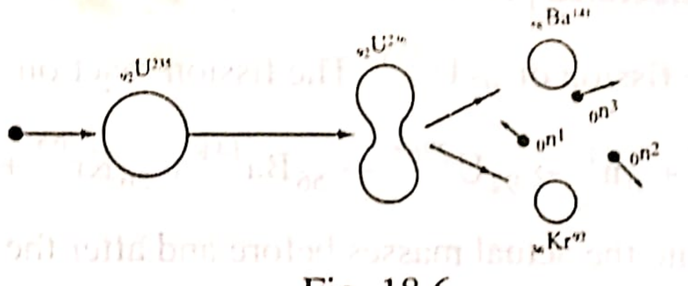


Fig. 18.6

As already mentioned, the fission products X and Y do not have unique values of mass number and atomic number. The fission products of $_{92}U^{235}$ fall into two groups.

(i) heavy group, with mass number lying between 130 and 155.

(ii) Light group, with mass number lying between 80 and 110.

The distribution of mass numbers in the fragments from the fission of $_{92}U^{235}$ is shown in Fig. 18.7 in which the percentage yields of the different products are plotted against mass number.

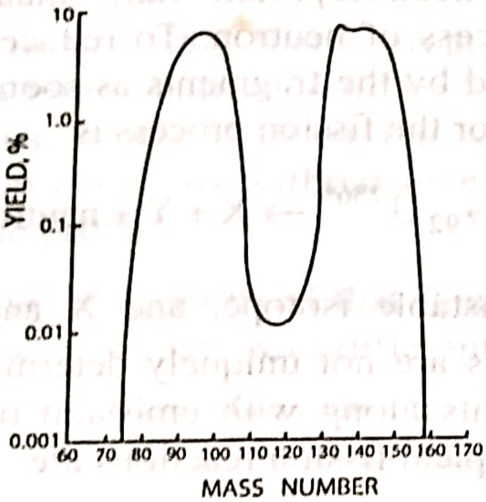


Fig. 18.7

Energy released in fission

A striking aspect of nuclear fission is the magnitude of the energy given off. This energy is produced because the original mass of the nucleus is greater than the sum of the masses of the fission products. The 'missing' mass is converted into energy according to Einstein's mass energy relation, $E = mc^2$.

The energy liberated per fission can be calculated as follows:

Consider the fission of $_{92}U^{235}$. The fission reaction is

$$_{92}U^{235} + {_0n^1} \rightarrow {_{92}U^{236*}} \rightarrow {_{56}Ba^{141}} + {_{36}Kr^{92}} + 3{_0n^1} + Q.$$

Let us estimate the actual masses before and after the fission,

mass of $_{92}U^{235}$	= 235.045733 amu
mass of $_0n^1$	= 1.008665 amu
Total initial mass (before fission)	= 236.054398 amu
mass of $_{56}Ba^{141}$	= 140.9177 amu
mass of $_{36}Kr^{92}$	= 91.8854 amu
mass of 3 neutrons (3 × 1.008665)	= 3.025995 amu
Total mass after fission	= 235.829295 amu

Decrease in mass $= 236.054398 - 235.829095$

$= 0.225303 = 0.2253$ amu

This decrease in mass is converted into energy. Since 1 amu = 931 MeV, the energy released in a fission process $= 0.2253 \times 931 = 209.8$ MeV.

Thus an energy in the neighbourhood of 200 MeV is released in a single nuclear fission, a remarkable figure for a single atomic event; chemical reactions liberate only a few electron volts per individual event. Most of the energy released in fission goes into the kinetic energy of the fission fragments. In the case of fission of U^{235}, about 83 per cent of the energy appears as kinetic energy of the fragments, about 2.5 percent as kinetic energy of the neutrons, and about 3.5 percent in the form of instantly emitted gamma rays. The remaining 11 percent is given off in the subsequent beta and gamma decays of the fission fragments.

A rough distribution of the energy released per fission is shown below:

	MeV	pJ
K.E. of fission products	168 MeV	26.9
K.E. of fast fission neutrons	5 MeV	0.8
β^- decay energy from fission products	4.8 MeV	0.768
Neutrino energy from β-decays	10.0 MeV	1.6
Immediate γ-ray energy	4.6 MeV	0.736
Fission produce γ-ray energy	6.9 MeV	1.104
Total	$\cong$ 199.3 MeV	$\cong$ 32

18.11 Chain reaction

Almost immediately after the discovery of nuclear fission it was recognized that because a neutron can induce fission in a suitable nucleus, with the consequent evolution of additional neutrons, a self-sustaining

sequence of fission should be possible (Fig. 18.8). Let us consider 1kg of $_{92}U^{235}$ which contains about $(6.023\times10^{26})/235 \cong 25\times10^{23}$ atoms.

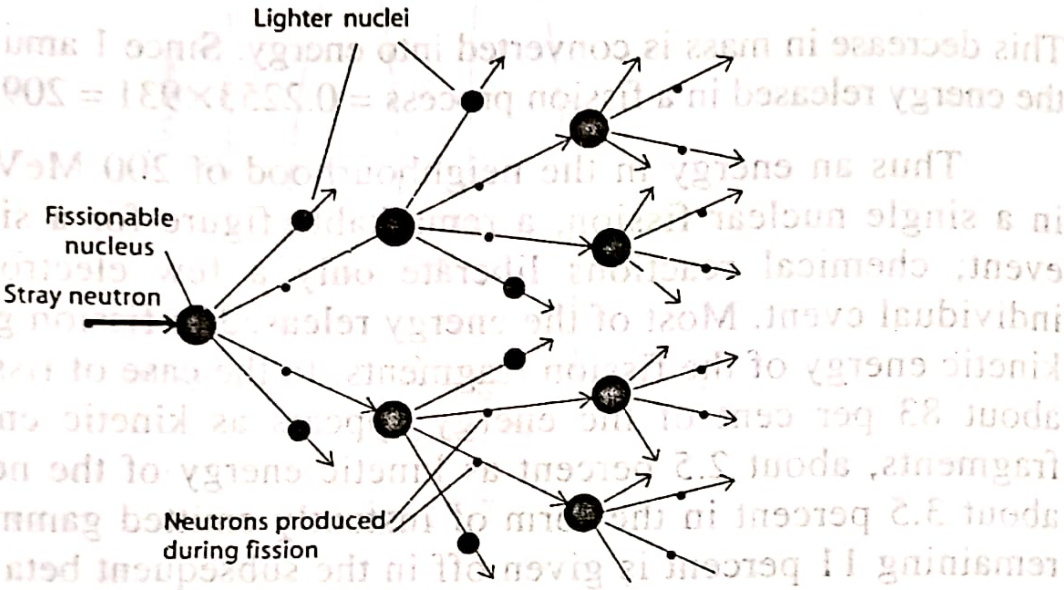


Fig. 18.8

Suppose a stray neutron causes fission in a uranium nucleus. Each fission will release on the average 2.5 neutrons. The velocity of neutrons among the uranium atoms is such that a fission capture of a thermal neutron by a U^{235} nucleus takes place in about 10^{-8}s. This fission, in turn, will release 2.5 neutrons. Assuming that all neutrons released in fission are available for inducing further fission reactions, the number of stages, n, of fission captures required to disrupt the entire mass of 1kg of U^{235} is

$$(2.5)^n = 25\times10^{23}; \quad \text{or} \quad n \cong 60$$

The time required for sixty fissions to take place

$$60\times10^{-8}\text{s} = 0.6\times10^{-6}\text{s} = 0.6\mu\text{s}.$$

An amount of energy of about 200 MeV is released in each fission. Therefore, a total of $200\times25\times10^{23} = 5\times10^{26}$ MeV of energy will be released in 0.6μs. The release of such enormous amount of energy in so short a time interval leads to a violent explosion. This results in powerful air blasts and high temperature of the order of 10^7K or more, besides intense radioactivity. This self-sustaining fission process described here is called *chain reaction*. The condition for such a chain reaction to occur in an assembly of fissionable material (*e.g.* uranium nuclei) is simple: at least one neutron produced during each fission must, on the average, initiate another fission. If too few neutrons initiate fissions, the reaction

will slow down and stop; if precisely one neutron per fission causes another fission, energy will be released at a constant speed. This is *controlled chain reaction.* Such a controlled chain reaction is used in *nuclear reactors.* In the other type of chain reaction the frequency of fissions is allowed to increase, the energy release will be so rapid that an explosion will occur. This type of reaction takes place in *atom bombs.*

Multiplication factor (k)

The ratio of secondary neutrons produced to the original neutrons is called the multiplication factor. It is given by

$$k = \frac{\text{Number of neutrons in any one generation}}{\text{Number of neutrons in the preceding generation}}$$

The fission chain reaction will be *critical* or steady when $k = 1$, it will be building up or *supercritical* when $k>1$ and it will be *subcritical* or dying down when $k<1$.

Critical size for maintenance of chain reaction

Consider a system consisting of uranium (as fissile material) and a material for slowing down neutrons, called *moderator*. Although, on an average 2.5 neutrons is released per fission, not all of them are available for future fission. The maintenance of the chain reaction depends upon a favourable balance of neutrons among the three processes given below:

(1) The fission of uranium nuclei which produces more neutrons than the number of neutrons used to induce fission.

(2) Non-fission processes, such as radiative capture of neutrons by the uranium and the parasitic capture by the different substance in the system and by impurities.

(3) Escape or leakage of neutrons through the surface of the system.

If the loss of neutrons due to the last two processes is less than the surplus produced in the first, a chain reaction takes place. Otherwise, it cannot take place.

The escape of neutrons takes place from the *surface* of the system while fission occurs throughout its *volume*.

As the area of a spherical mass is $4\pi r^2$, the escape rate varies as r^2 and the production rate varies as r^3 $\left(\because \text{volume} = \frac{4}{3}\pi r^3\right)$.

$$\therefore \quad \frac{\text{escape rate}}{\text{production rate}} \propto \frac{1}{r}$$

Thus, the greater the size of the system, the lesser will be the probability of the escape of neutrons. It will therefore be possible to reduce the probability of escape of neutrons by increasing the volume of the system. In this case, the production of neutrons will be more than the loss due to other causes and a chain reaction can be maintained. Thus there is a critical size for the system. *Critical size of a system containing fission material is defined as the minimum size for the number of neutrons produced in the fission process just balance those lost due to leakage and non-fission capture.* The mass of the fissionable material in this critical size is called *critical mass.* If the size is less than the critical size, a chain reaction is not possible.

Natural uranium and chain reaction

Natural uranium contains only 0.72 percent of the fissionable isotope U^{235}. The more abundant (99.28 percent) U^{238} isotope readily captures fast neutrons but does not undergo fission as a result. However, U^{238} has only a small cross section for the capture of *slow* neutrons, whereas U^{235} has a very high cross section for fission capture of slow neutrons; it undergoes fission even by thermal neutrons (energy 0.038 eV). Thus if the fast neutrons that are liberated in fission can be rapidly slowed down, this will prevent the unproductive absorption of the neutrons by U^{238} and at the same time promote further fission in U^{235}.

The neutrons can be slowed down by distributing among lumps or rods of uranium a material called *moderator.* The essential function of the moderator is to slow down the fast neutrons of fission very rapidly by elastic collisions to thermal energies, so that the chance of non-fission capture of U^{238} becomes small, and consequently the fission of U^{235} increases as illustrated in (Fig.18.9). The moderator must necessarily possess the property of itself not

capturing neutrons readily, since otherwise it will augment, instead of diminishing, the collapse of the chain reaction.

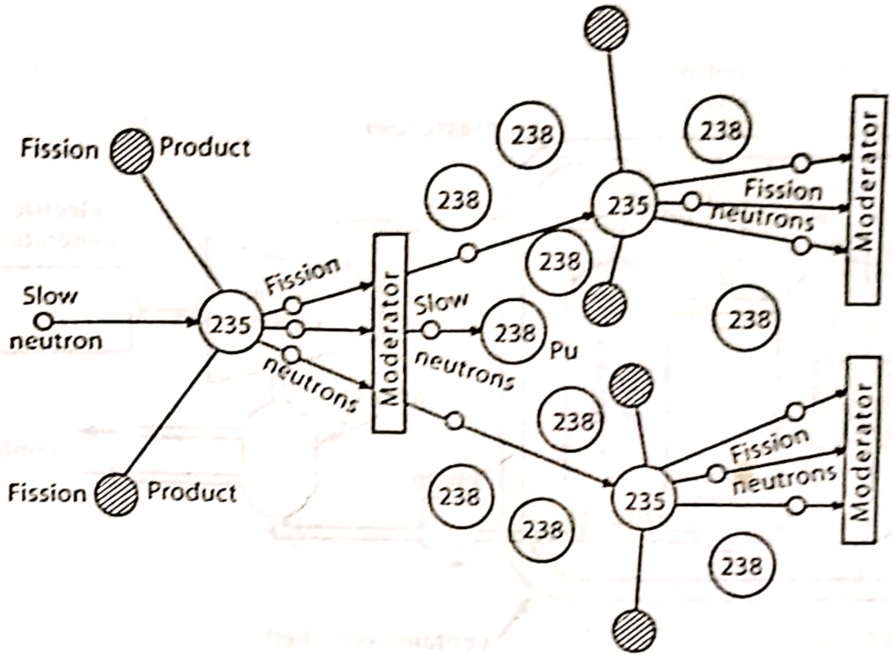


Fig. 18.9

While the exact amount of energy lost by a moving body that collides elastically with another depends on the details of the interaction, in general the energy transfer is maximum when the participants are of equal mass. The greater the difference between the masses, the greater the number of collisions needed to slow a neutron down, and the longer the period in which it is in danger of being captured by a U^{238} nucleus. The moderators in common use are ordinary (light) water, 'heavy' water whose molecules contain deuterium atoms instead of ordinary hydrogen atoms, and graphite, a form of pure carbon.

18.12 Nuclear reactors

A *nuclear reactor* is a device in which a chain reaction involving nuclear fission can be initiated and controlled. A reactor is a very efficient source of energy: the fission of 1gm of a suitable fissionable material per day evolves energy at the rate of about 1MW (10^6 watt). For producing energy at the same rate combustion of 2.6 tons of coals per day will be needed. The energy liberated in a nuclear reactor becomes heat in its interior, and this heat is removed by circulating a liquid or gas coolant. The hot coolant is then used to boil water, and the resulting

2009-2010

steam is fed to a turbine that can power an electric generator to produce electricity. The basic design of a power reactor is shown in Fig. 18.10.

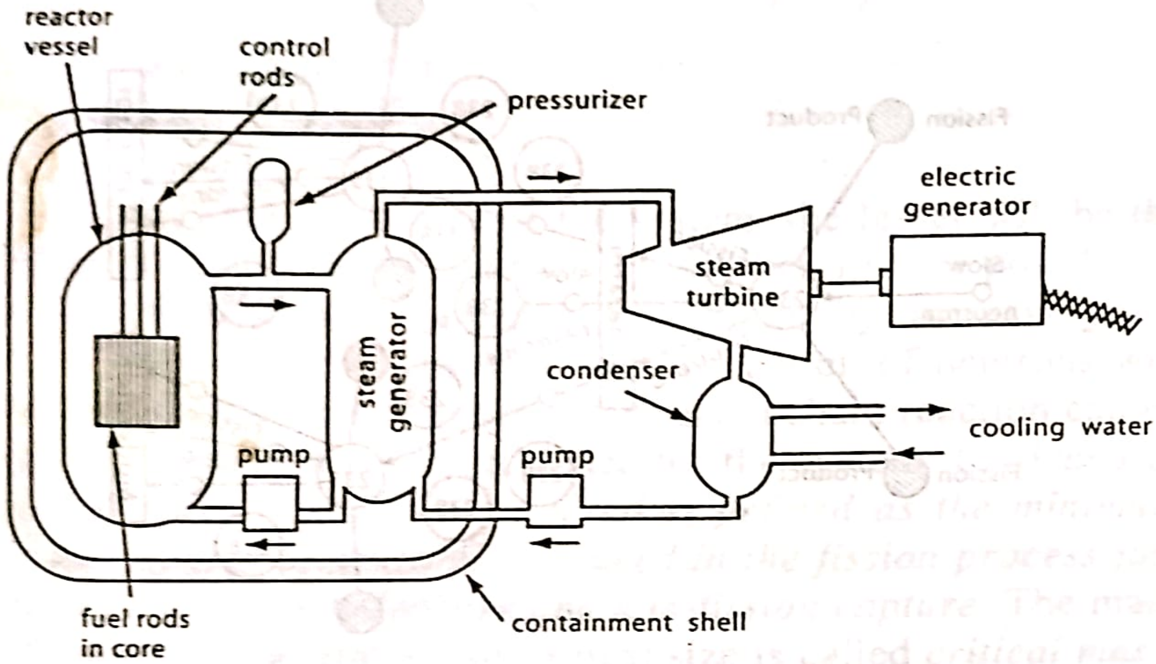


Fig. 18.10

The essential elements of a nuclear reactor are

(1) The fissionable material called fuel

(2) Moderator

(3) Neutron reflector

(4) Cooling system, and

(5) The safety and control system.

(1) **The fissionable substance.** The commonly used fissionable materials are the isotopes of uranium (U^{235} and U^{238}), the thorium isotope Th^{232} and the plutonium isotopes Pu^{239}, Pu^{240} and Pu^{241}.

(2) **Moderator.** The function of a moderator is to slow down the highly energetic fission neutrons (fast neutrons) to thermal energies. Commonly used 'moderator' materials are graphite, heave water (D_2O), beryllium, etc. Ideal moderating materials should have low atomic weight and low absorption cross-sections for neutrons.

(3) **Neutron reflector.** One of the important problems in the design of a reactor is the loss of neutrons through the reactor surface. Each fission in U^{235} releases an average of 2.5 neutrons, so no more than 1.5 neutrons per fission can be lost if a self-sustaining chain reaction is to occur. This can be solved by increasing the reactor size, since a large object has less surface area in proportion to its volume than a small one. Leakage of neutrons through the surface can be very much reduced and the neutron flux in the interior can be increased by the use of reflector on the surface of reactors. Good reflector materials normally have high scattering cross-section and low-absorption cross-section for neutrons.

(4) **Cooling system.** As the fission fragments are slowed down in the fissionable substance and the moderator, a tremendous amount of heat is evolved. The coolant or heat-transfer agent is pumped through the reactor core. Then, through a heat-exchanger, water is converted into steam. The steam produced runs conventional turbines to produce electricity. Cooling agents normally used are water, steam, He, Co_2, air and certain molten metals and alloys.

(5) **Control and safety system.** The control system enables the chain reaction to be controlled and prevent it from spontaneously running away. The time between the release of a fission neutron and its later absorption is less than a millisecond. In order to control the rate of the chain reaction, the reactor incorporates rods made of material such as cadmium or boron that readily absorbs slow neutrons. As these rods are inserted further and further into the reactor, the reaction rate is progressively damped and the reactor can be made to *die down*. On the otherhand, by pulling the rods out, the reactor can be made to *build up*. The safety systems protect the space surrounding the reactor against intensive neutron flux and gamma rays existing in the reactor core. This is achieved by surrounding the reactor with massive walls of concrete and lead which would absorb neutrons and gamma rays.

18.13 The Atom Bomb

The principle of uncontrolled nuclear fission is used to construct the atom bomb. The atom bomb essentially consists of two pieces of U^{235} (or Pu^{239}) each smaller than the critical size and source of neutrons. The two subcritical masses of U^{235} in the form of hemispheres are kept apart by using a separator aperture (Fig.18.11). When the bomb has to be exploded, a third well fitting cylinder of U^{235}, whose mass is also less than the critical mass, is propelled so that it fits in or fuse together with the other two pieces. The total quantity of U^{235} is now greater than the critical mass. A neutron source is activated and an uncontrolled chain reaction takes place resulting in a terrific explosion.

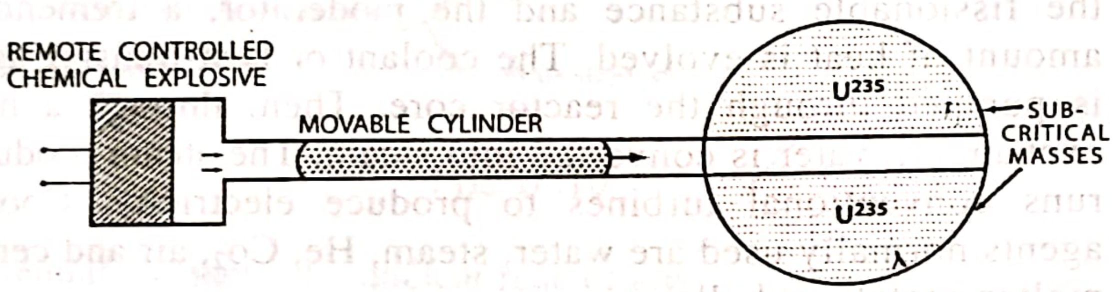


Fig. 18.11

The explosion of an atom bomb releases a tremendously large quantity of energy in the form of heat, light and radiation in a very short time. A temperature of millions of degrees and a pressure of millions of atmospheres are produced. Such explosions produce shock waves. They are very dangerous because the waves spread radioactivity in air and cause loss of life. The release of dangerously radioactive γ-rays, neutrons and radioactive materials presents a health hazard over the surroundings for a long time. The radioactive fragments and isotopes formed out of explosion adhere to dust particles thrown into space and then fall back to earth, even at very distant places. This is termed as radiation *fall-out.*

Atom bombs were used in world war II and exploded over Nagasaki and Hiroshima in Japan. Each bombs released energy equivalent of 20,000 tons of TNT.

18.14 Nuclear Fusion

The distance of the earth from sun is 150 million km. A surface $1 m^2$ in area exposed to the vertical rays of the sun receives energy at the rate of about 1.4 kW. This corresponds to an energy of 4×10^{26}W (3.8×10^{26} joules) that is being radiated by the sun per second. And the sun has been emitting energy at this rate for billion of years. Where does it come from?

The basic energy-producing process in the sun is a process called *nuclear fusion*. In this process, small nuclei fuse together to form a single heavy nucleus. The mass of the single nucleus formed is less than the mass of the 'fused' nuclei. This difference in mass is converted into energy according to Einstein's mass energy relation, $E = mc^2$.

The fusion of nuclei can take place in several different reaction sequences, the most common of which, the *proton-proton* cycle as suggested by Bethe.

Proton-proton cycle. The first step in the proton-proton cycle is the formation of deuterons by the direct combination of two protons, which is accompanied by the emission of a positron and a neutrino:

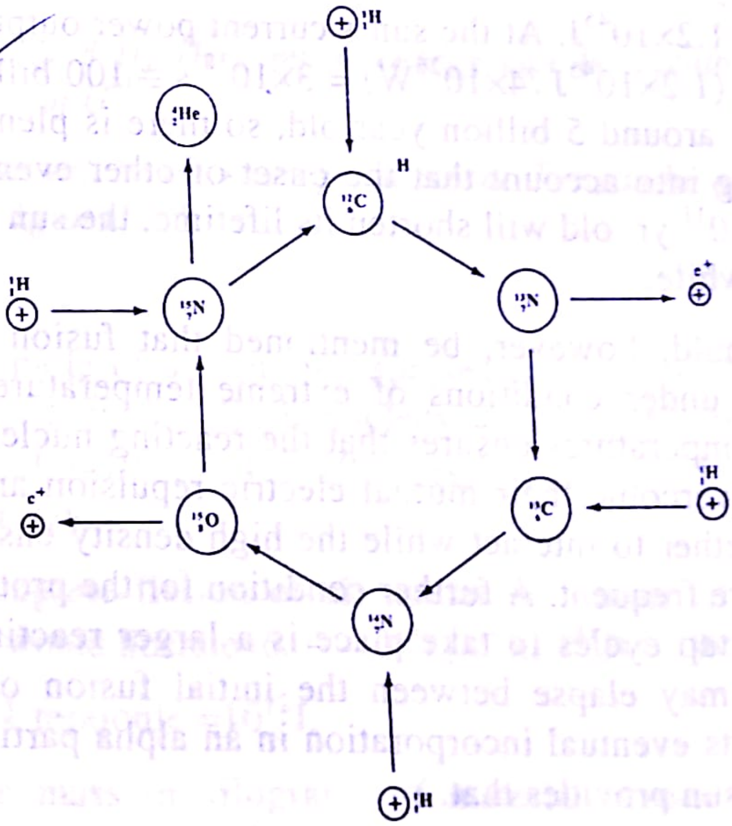


Fig. 18.12

$${}_1H^1 + {}_1H^1 \rightarrow {}_1H^2 + e^+ + \upsilon$$

A deuteron may then join with a proton to form a ${}_2He^3$ nucleus;

$${}_1H^2 + {}_1H^1 \rightarrow {}_2He^3 + \gamma$$

Finally two ${}_2He^3$ nuclei react to produce a ${}_2He^4$ nucleus plus two protons

$${}_2He^3 + {}_2He^3 \rightarrow {}_2He^4 + {}_1H^1 + {}_1H^1$$

The total energy evolved in the sequence is $(\Delta m)c^2$, where Δm is the difference between the mass of four protons and the mass of an alpha particle plus two positrons; it turns out to be 24.7 MeV.

Since 1 MeV = 1.6×10^{-13}J; 24.7 MeV = 4×10^{-12}J, and the sun's power output of 4×10^{26}W means the above sequence of reactions must take place 10^{38} times per second. The sun's mass of 2×10^{30} kg corresponds to the mass of about 1.2×10^{57} protons. Four of these protons are used up in each of the above sequence. Therefore, the number of possible sequences that can occur is about 3×10^{56}. The total energy that will be released in these sequences is $(4\times10^{-12}$ J) $(3\times10^{56}) = 1.2\times10^{45}$J. At the sun's current power output this means a lifetime of $(1.2\times10^{45}\text{J} / 4\times10^{26}\text{W}) = 3\times10^{18}$ s ≡ 100 billion years. The sun is only around 5 billion year old, so there is plenty of fuel left. Even taking into account that the onset of other events long before the sun is 10^{11} yr. old will shorten its lifetime, the sun will still shine for a long while.

It should, however, be mentioned that fusion reactions can occur only under conditions of extreme temperature and density. The high temperatures ensures that the reacting nuclei have enough energy to overcome their mutual electric repulsion and come close enough together to interact while the high density ensures that such collisions are frequent. A further condition for the proton-proton and other multistep cycles to take place is a larger reacting mass since much time may elapse between the initial fusion of a particular proton and its eventual incorporation in an alpha particle. The large mass of the sun provides that.

The temperature of the core of the sun is believed to be about 1.5×10^7 K, which is sufficient for the proton-proton cycle to occur

there. The same is true for other stars. Some of these stars may have hotter interiors, and in them the *carbon-nitrogen* cycle predominales

Carbon-Nitrogen Cycle. The cycle proceeds in the following way(Fig.18.12).

$$C_6^{12} + {}_1H^1 \rightarrow {}_7N^{13*} + \gamma$$

$${}_7N^{13*} \rightarrow C_6^{13} + e^+ + \upsilon$$

$$C_6^{13} + {}_1H^1 \rightarrow {}_7N^{14} + \gamma$$

$${}_7N^{14} + {}_1H^1 \rightarrow {}_8O^{15*} + \gamma$$

$${}_8O^{15*} \rightarrow {}_7N^{15} + e^+ + \upsilon$$

$${}_7N^{15} + {}_1H^1 \rightarrow C_6^{12} + {}_2He^4$$

The net result is again the formation of an alpha particle and two positrons from four protons, with the evolution of about 24.7 MeV. C_6^{12} acts as a catalyst for the cycle since it reappears at the end of the cycle.

Example 18.8 (i) How much energy can be obtained from the fission of 1 mole of U^{235}?

(ii) What will be the energy and power that can be obtained from the fission of 1 kg of U^{235}?

Soln.

(i) One mole of any atomic species contains 6.02×10^{23} atoms (the Avogadro number). Each nucleus undergoing fission releases 200 MeV or $(200\times10^6$ ev$)(1.6\times10^{-19}$J/ev$) = 3.2\times10^{-11}$ J or 32×10^{-12}J or 32 pJ of energy.

Therefore complete fission of all nuclei in 1 mole of U^{235} or any other element would liberate $(6.02\times10^{23})\ (3.2\times10^{-11}\text{J}) = 19\times10^{12}\text{J} = 19$ TJ

where 1TJ = 1 terajoule $=10^{12}$J.

(ii) The mass in kilograms of material in one mole of any element is equal to 10^{-3} times the conventional atomic weight. Therefore the mass in kg of one mole of U^{235}

$= 235 \times 10^{-3}$ kg.

Hence complete fission of 235×10^{-3} kg will liberate 19 TJ of energy.

Therefore complete fission of 1kg of U^{235} liberates

$$\frac{19}{235\times10^{-3}} \text{TJ} = 81 \text{ TJ of energy}$$

Now 1J = 1 watt-second (W.s)

$\therefore$ 1 TJ = 10^{12} W.s = 11.6 megawatt days (MW.d)

Hence, the complete fission of 1kg of U^{235} will produce power equivalent to (11.6×81) MW.d = 939.6 MW.d

Example 18.9 A nuclear reactor using $_{92}U^{235}$ *is to operate at a power level of 250 megawatts. If the energy released per fission of* U^{235} *is 200 MeV, calculate the rate of consumption of* U^{235} *per year. Assumè that there are no losses and that the mass of an atom is equal to the sum of the masses of the nucleons in the nucleus where each nucleon has a mass equal to 1 a.m.u.* $(1.6\times10^{-27}$ *kg).*

Soln.

Output power = 250×10^6 watts = 2.5×10^8 J/s.

Energy released per fission = 200 MeV

$= 200 \times10^6$ ev $= 200\times10^6 \times1.69\times10^{-19}$ J

$= 3.2\times10^{-11}$ J.

Number of fission per second

$$= \frac{2.5\times10^8}{3.2\times10^{-11}} = 7.81\times10^{18}$$

Mass of each atom of $_{92}U^{235}$

$= 235\times1.6\times10^{-27}$ kg (1 a.m.u. = 1.6×10^{-27} kg)

$= 3.76 \times10^{-25}$ kg.

Consumption per second

$= 7.81\times10^{18}\times3.76\times10^{-25}$ kg

$= 2.936\times10^{-6}$ kg.

Consumption per year

$= 2.936\times10^{-6}\times86400\times365$

$= 92.59$ kg.

Example 18.10 *A deuterium reaction that occurs in experimental fusion reactor is H^2 (d, p) H^3 followed by H^3 (d, n) He^4. (a) Compute the energy release in each of these. (b) Compute the total energy release per gram of the deuteron used in the fusion. (c) Compute the percentage of rest mass of deuteron released as energy. (d) Compare U^{235} fission with deuteron fusion as a source of energy release. Given H^2 = 2.014740 amu, H^3 =3.17005 amu, U^{235} = 235.1178 amu.*

Soln.

(a) (i) The fusion reaction H^2 (d, p) H^3 is

$$_1H^2 + {}_1H^2 \rightarrow {}_1H^3 + {}_1H^1 + Q$$

Decrease in mass in the reaction, Δm

$= 2.014740 + 2.014740 - 3.017005 - 1.008145$

$= 0.004330$ amu.

$\therefore$ Energy release $= 0.004330 \times 931$ MeV

$= 4.031$ MeV.

(b) (ii) The reaction H^3 (d, n) He^4 is

$$_1H^3 + {}_1H^2 \rightarrow {}_2He^4 + {}_0n^1 + Q$$

$\Delta m = 3.017005 + 2.014740 - 4.003179 - 1.008986$

$= 0.01958$

Energy released $= 0.01958 \times 931$ MeV $= 18.23$ MeV

$\therefore$ Total energy released $= (4.031 + 18.23)$ MeV

$= 22.26$ MeV.

(b) The total energy released in (a) is from the fusion of 3 deuterium ($_1H^2$) nuclei.

$\therefore$ Energy released per $_1H^2$ nuclei

$= 22.26 \div 3 = 7.42$ MeV.

Number of nuclei in 1 gm of $_1H^2$

$$= \frac{6.02 \times 10^{23}}{2.01474}$$

$\therefore$ Total energy release from 1 gm of $_1H^2$

$$= \frac{6.02 \times 10^{23}}{2.01474} \times 7.42 = 2.217 \times 10^{24} \text{ MeV.}$$

(c) Energy equivalent of one H^2 nucleus

$= 2.01474 \times 931$ MeV

Average release of energy per H^2 nucleus $= 7.42$ MeV

$\therefore$ The percentage of the rest mass of deuteron released as energy

$$\frac{7.42}{2.01474 \times 931} \times 100 = 0.3956\%$$

(d) In U^{235}, 200 MeV is released per fission of U^{235} nucleus.

$\therefore$ Percentage of the rest mass of U^{235} released as energy

$$\frac{200}{235.1 \times 931} \times 100 = 0.09317\%$$

$$\therefore \quad \frac{\text{Engery release from } H^2 \text{ fusion}}{\text{Engery release from } U^{235} \text{ fission}} = \frac{0.3956\%}{0.09317\%} = 4.247$$

***Example 18.11** A city requires on the average 100 megawatts of electrical power per day and this is to be supplied by a nuclear reactor of efficiency 20%. Using U^{235} as a nuclear fuel, calculate the amount of fuel required for one day's operation. Energy released per fission of U^{235} nuclide is 200 MeV.*

Soln.

Energy required per day

$$E_1 = 100\times10^6 \times 86400 \text{ J}$$

$$= 8.64 \times (10^{12} \text{ J})$$

Suppose amount of fuel required = m kg

No. of atoms in m kg of U^{235}

$$= \frac{6.023\times10^{26}}{234}\times m$$

Energy released per fission = 200 MeV.

$$= 200 \times 10^6 \times 1.6 \times 10^{-19} \text{J}$$

$$= 3.2 \times 10^{-11} \text{J}$$

∴ Total energy released

$$= \left(\frac{m\times6.023\times10^{26}}{235}\right)\times\left(3.2\times10^{-11}\right) \text{J}.$$

Efficiency = 20%

$$\therefore \quad \text{Useful energy, } E_2 = \left[\frac{0.20\times m\times6.023\times10^{26}\times3.2\times10^{-11}}{235}\right] \text{J}$$

$$= \left[16.403\times10^{12} \text{ m}\right] \text{J}$$

For E_1 to be equal to E_2,

$$8.64\times10^{12} \text{ J} = 16.403\times10^{12} \text{ m J}$$

$$\therefore \quad m = \frac{8.64\times10^{12}}{16.403\times10^{12}} = 0.5267 \text{ kg/day.}$$

EXERCISES

1. Give a short account of the (i) charge (ii) mass (iii) radius (iv) binding energy (v) magnetic moment and (vi) quadrupole moment of the nucleus.

2. Discuss the basic properties of an atomic nucleus.

3. Define 'mass defect', 'packing fraction' and 'binding energy' of a nucleus.

4. Sketch the curve of the binding energy per nucleon for the most stable nucleus against the corresponding mass number and discuss its nature. What information can you obtain from this curve?

5. What do you understand by atomic mass unit and the binding energy of the nucleus? Derive the relation between atomic mass unit and electron volt.

6. Briefly describe the nature of the nuclear forces.

7. Describe the 'liquid drop model' of the nucleus. How can the semi-empirical mass formula be derived from it?

8. Write down Weizsacker's semi-empirical mass formula and explain each term in the formula.

9. Explain the main features of the 'nuclear shell model'. What are magic numbers?

10. What is nuclear fission? What is the source of release of energy in nuclear fission?

11. Outline the process of nuclear fission and its important application.

12. What is nuclear fission? Calculate the energy released in a fission of U^{235} nucleus. How is this energy distributed?

13. Describe briefly the various components of a nuclear reactor. Explain the principle of its working. When is the reactor said to be critical?

14. What is nuclear fusion? Briefly describe the proto-proton cycle and the carbon-nitrogen cycle of the production of stellar energy. Distinguish between fission and fusion.

15. Show that the density of nuclear matter is independent of mass number. Obtain an estimate of the nuclear density.

16. Given the following isotope masses:

    ${}_3Li^7 = 7.016004$, ${}_3Li^6 = 6.015125$ and ${}_0n^1 = 1.008665$. Calculate the B.E. of a neutron in the ${}_3Li^7$ nucleus. Express the result in MeV and joule.
    [7.5 MeV, 1.18×10^{-12}J]

17. Calculate the binding energies of the following isobars and their binding energy per nucleon. $_{28}Ni^{64}$ =63.9279 a.m.u; $_{29}Cu^{64}$ = 63.9297 a.m.u; $_{30}Zn^{64}$ = 63.9297 a.m.u. Given m_p = 1.007825 amu and m_n = 1.008665 amu
[561.5 MeV, 8.774 MeV, 558.6 MeV, 8.73 MeV, 558.3 MeV, 8.72 MeV]

18. Find the release of energy when two $_1H^2$ nuclei fuse together to form $_2He^4$ nucleus. The binding energy per nucleon of H^2 and He^4 is 1.1 MeV and 7.0 MeV respectively.

[B.E. for $_2He^4$ = 28 MeV; B.E. for $_1H^2$ = 2.2 MeV.
Δm = 2[Mass of $_1H^2$] – [Mass of $_2He^4$]
= 2[$m_p + m_n - 2.2$] – [2 m_p + 2 m_n–28]
= 2 m_p + 2 m_n – 4.4 - 2 m_p + 2 m_n – 28
= 23.6 MeV]

19. Derive a formula for the atomic number of the most stable isobar of a given A and use it to find the most stable isobar of A = 25.

$$[E_b = a_1A - a_2A^{2/3} - a_3\frac{Z(Z-1)}{A^{1/3}} - a_4\frac{(A-2Z)^2}{A}(\pm, 0)\frac{a_5}{A^{3/4}}$$

For E_b to be maximum, $dE_b/dz = 0$

$$\frac{dE_b}{dz} = -\frac{a_3}{A^{1/3}}(2Z-1) + \frac{4a_4}{A}(A-2z) = 0$$

$$Z = \frac{a_3A^{-1/3} + 4a_4}{2a_3A^{-1/3} + 8a_4A^{-1}} = \frac{0.6A^{-1/3} + 76}{1.2A^{-1/3} + 152A^{-1}}$$

For A =25, Z = 11.7 for which we conclude Z = 12 should be the atomic number of most stable isobar of A = 25. This nuclide is $_{12}Mg^{25}$ which is the only stable isobar with A = 25. The other isobar $_{11}Na^{25}$ and $_{13}Al^{25}$ are both radioactive].

20. A reactor is developing energy at the rate of 3000 kW. How many atoms of U^{235} undergo fission per second? How many kilograms of U^{235} would be used in 1000 hours of operation assuming that on an average energy of 200 MeV is released per fission. **[9.4×10^{16}; 0.1321 kg]**

21. Assuming the energy released per fission of U^{235} is 200 MeV, calculate the rate in kg per year at which fission should occur in a nuclear reactor in order to operate at a power level of 1 watt. **[1.636×10^{-19} kg]**

22. When an atom of U^{235} undergoes fission in a reactor 200 MeV energy is liberated. Suppose that the power output is 800 MW and the reactor is 25% efficient. (i) How many uranium atom does it consume in one day, (ii) What mass of uranium does it consume in one day? **[9.64 × 10^{26} atoms; 376.31 kg]**

23. The fusion reaction $2\,{}_1H^2 \rightarrow {}_2He^4$ + energy, is proposed to be used for the production of industrial power. Assuming the efficiency of the process to be 30%, find how many kg of deuterium will be consumed in a day for an output of 50,000 kW. Given: mass of ${}_1H^2$ = 2.01478 a.m.u; mass of ${}_2He^4$ = 4.00388 a.m.u.

[Δm = 0.02568 a.m.u; Energy released = 0.02568 × 931 MeV = 23.91 MeV. Since the efficiency of the process is 30%, the actual output is 7.173 MeV. ∴ Actual output per deuterium atom = 7.173/2 = 3.587 MeV = $(3.587 \times 1.6 \times 10^{-13})$J = 5.74×10^{-13} J. The required output is 5000 kW = 5×10^7 J. ∴ No of deuterium atoms required to be consumed to produce an output of 5×10^7 J = $(5\times10^7)/(5.74\times10^{-13})$ = 8.71×10^{19}. The mass of 8.71×10^{19} deuterium atoms = $\frac{2\times8.71\times10^{19}}{6.023\times10^{26}}$ = 2.90×10^{-7} kg.

This is the consumption in one second.

∴ The consumption of deuterium in one day = $2.90 \times 60 \times 60 \times 24 \times 10^{-7}$ = 2.51×10^{-2} kg.]

CHAPTER XIX

SOLID STATE PHYSICS

19.1 Crystals and their structures

A crystal is just a solid having uniform chemical composition and some recognizable features which are the same for all specimens. For example, the crystals are bounded by plane surfaccs which intersect at definite angles. Though two crystals of the same substance may look different in external appearance, the angles between the corresponding faces are always the same. Such a definite pattern of these angles can be expected only on the assumption that the atoms or molecules are arranged in a regular array. It is this geometrical perfection of a crystal which distinguishes it from a non-crystalline (or amorphous) substance like glass which splits into pieces of an infinite variety of different shapes, marked by edges and surfaces that are seldom plane. Thus, it is natural to conclude that whereas the atoms or molecules of a crystal are arranged one next to the other strictly according to a rigorous geometrical scheme, the atoms of an amorphous substance are irregularly arranged, like sand grains in a sand pile. Crystals can be further distinguished from non-crystalline (amorphous) substances by their property of anisotropy, *i.e.*, the dependence of quantities like linear compressibility, temperature expansion co-efficient, refractive index, electrical and heat conductivities and the velocity of sound in the direction through the crystal in which they are measured.

A crystal is composed of fundamental units called *basis* which may be simple atoms or molecules or a group of melocules, identical in composition, arrangement and orientation. Each atom or melocule is fixed at a *definite point in space at a definite distance from and in a definite angular orientation to* all others surrounding it. This fundamental unit is repeated an infinite number of times through the structure so that the position of one particular unit with respect to its neighbor is exactly the same as that of any other unit. One can replace each unit by a geometrical point. The result is a pattern of points having the same geometrical properties of the crystal. This *infinite array* of points in three dimensions in which every point has surroundings or environments identical to that of every other point in the array is called a *space lattice* or *crystal lattice* or simply *lattice*. The crystal lattice is distinguished from the crystal structure by the fact that a crystal structure is formed by associating with every lattice point a unit assembly or basis.

$$\text{lattice} + \text{basis} = \text{crystal structure}$$

19.2 Crystal lattice and unit cell

The entire lattice structure of a crystal is found to consist of identical blocks which are repeated again and again to form the lattice structure. The *smallest block* or the *geometric figure* which can generate the complete crystal by repeating itself in three dimensions is called the *unit cell.* The unit cell may be defined, in general, as that *volume of a solid* from which the entire crystal can be constructed by translational repetition in three dimensions (Fig. 19.1).

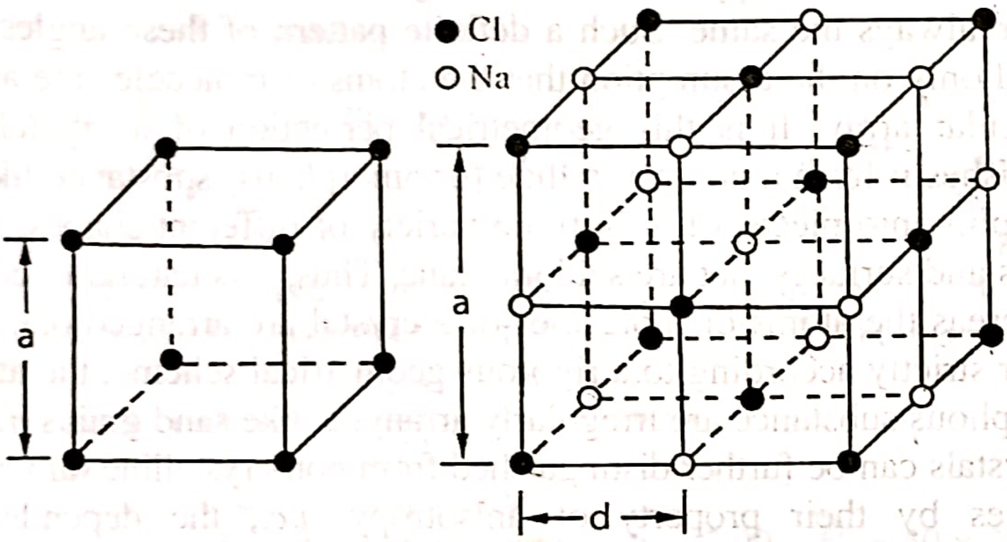


Fig. 19.1

In fact, a unit cell is the smallest part of a crystal having all the structural properties of the given lattice.

If each atom in a lattice is replaced by a point, then each such point is called a *lattice point*. The arrangement of these points is referred to as the (three - dimensional) lattice array. The basic interplaner distance or principal grating space, d, is shown for the cubic crystal KCl (Fig. 19.2). Note that d is the distance between adjacent atoms. The length of the side of a unit cell is the distance between the atoms of the same kind. This is equal to 2d for the cubic crystal in Fig. 19.2. The length of the side of a unit cell is equal to the basic distance d only in the case of simple cubic structures in which all of the atoms are of the same kind. Some pure metals form crystals of this type. This basic interatomic distance d can be

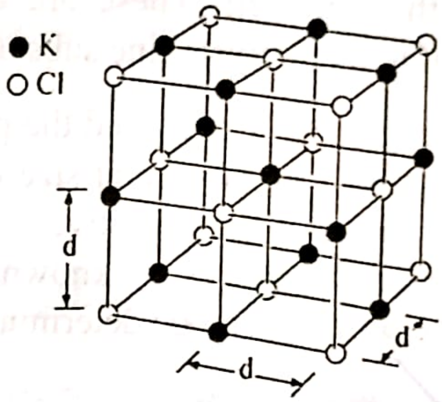


Fig. 19.2

computed from the knowledge of the molecular weight of the crystalline compound, Avogadro's number, the density of the material and its crystalline form.

19.3 Lattice parameters of a unit cell

The three sides of a unit cell are called the *crystallographic axes* and are so chosen that they have directions such that they pass regularly through the lattice points associated with like-atoms. An arbitrary arrangement of crystallographic axes, marked X, Y and Z, defining a unit cell is shown in Fig. 19.3. The intercepts a, b and c define the dimensions of a unit cell and are known as its *primitives* or *characteristic intercepts* on the axes. The angles α, β and γ between the three axes are called *interfacial angles*. The primitives and the interfacial angles constitute the

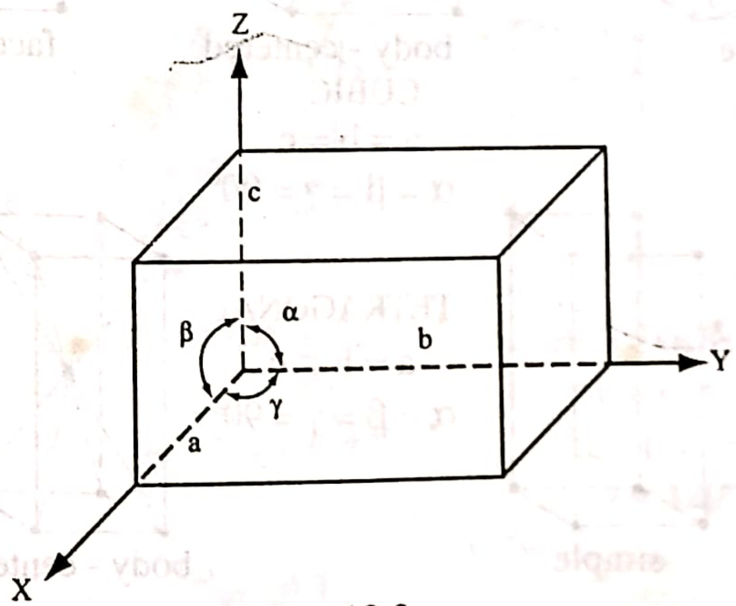


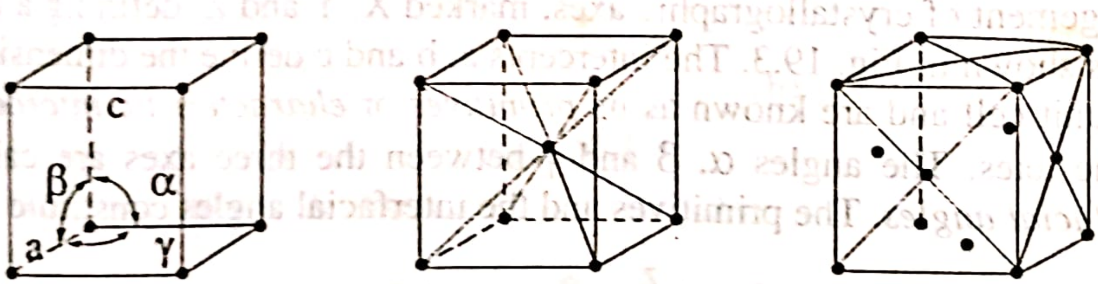
Fig. 19.3

lattice parameters of the unit cell. These are also referred to as the geometrical constants of a given crystalline substance.

If the interfacial angels α, β and γ and the primitives a, b and c are known, then the form as well as the actual size of the unit cell are also known. But if only the ratios of the primitives - not their actual values, and the values of the interfacial angles are known, then only the form and not the actual size of the unit cell can be determined.

19.4 Bravais Lattices

Unit cells fall into 14 categories, corresponding to the 14 possible different space lattices. In the simplest crystals there is a single atom at each lattice point, but often an assembly of two or more atoms occupies each lattice point, where each assembly is the same in composition, arrangement and orientation. Consequently a distinction should be made between the structure of a crystal, which is the actual ordering in space of its constituent atoms, ions or molecules, and the corresponding crystal lattice, which is a geometrical abstraction useful for classifying the crystal. According to *Bravais* who introduced the space lattice idea as early as 1880, there are only 14 possible types of space lattice. Fig. 19.4.

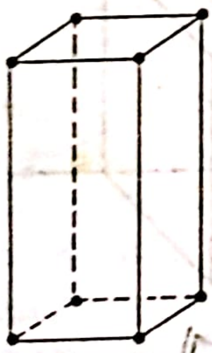


simple   body - centered   face - centered

CUBIC

$a = b = c$

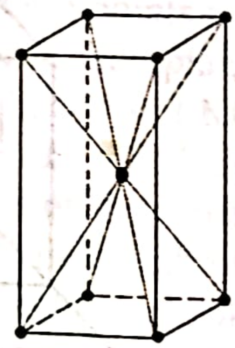
$\alpha = \beta = \gamma = 90^\circ$



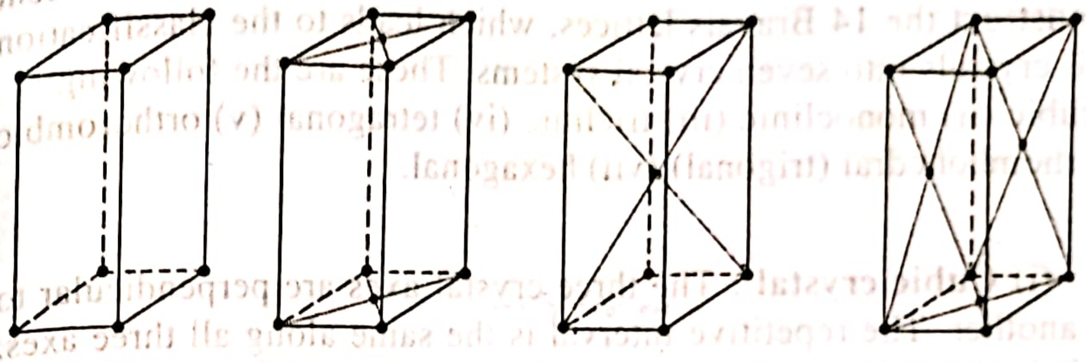
TETRAGONAL

$a = b \neq c$

$\alpha = \beta = \gamma = 90^\circ$



simple   body - centered

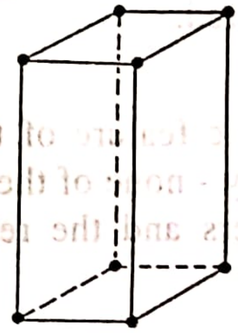


simple base - centered body - centered face - centered

ORTHORHOMBIC

$a \neq b \neq c$

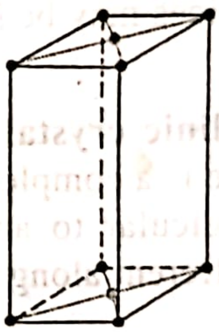
$\alpha = \beta = \gamma = 90°$



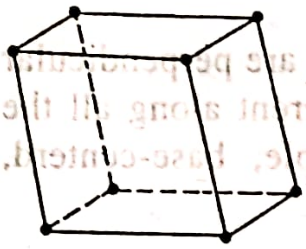
MONOCLINIC

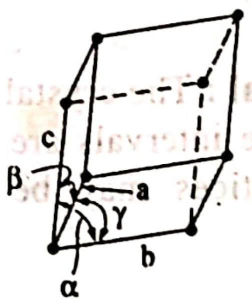
$a \neq b \neq c$

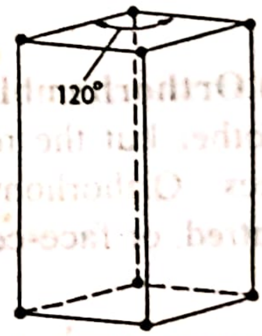
$\alpha = \beta = 90° \neq \gamma$



simple base - centered







TRICLINIC

$a \neq b \neq c$

$\alpha \neq \beta \neq \gamma$

TRIGONAL

$a = b = c$

$\alpha = \beta = \gamma \neq 90°$

HEXAGONAL

$a = b \neq c$

$\alpha = \beta = 90°$

$\gamma = 120°$

Fig. 19.4

From Fig. 19.4 it is clear that seven sets of axes are sufficient to construct the 14 Bravais lattices, which leads to the classification of all crystals into seven crystal systems. These are the following: (i) cubic (ii) monoclinic (iii) triclinic (iv) tetragonal (v) orthorombic (vi) rhombohedral (trigonal) (vii) hexagonal.

(i) Cubic crystal : The three crystal axes are perpendicular to one another. The repetitive interval is the same along all three axes. Cubic lattices may be simple, body-centred or face centred.

(ii) Monoclinic crystal : Two of the crystal axes are not perpendicular to each other, but the third is perpendicular to both of them. The repetitive intervals are different along all the three axes. Monoclinic lattices may be simple or base-centred.

(iii) Triclinic crystal : The characteristic feature of triclinic crystal structure is a complete lack of symmetry - none of the crystal axes is perpendicular to any one of the others and the repetitive intervals are different along all the three axes.

(iv) Tetragonal crystal : The crystal axes are perpendicular to one another. The repetitive intervals are same along two axes while that along the third axis is different. Tetragonal lattices may be simple or body-centred.

(v) Orthorhombic crystal : The crystal axes are perpendicular to each other but the repetitive intervals are different along all the three axes. Orthorhombic lattices may be simple, base-centerd, body-centred, or face-centred.

(vi) Rhombohedral (also known as trigonal) crystal : The angles between each pair of crystal axes are same but they are not 90°. The repetitive intervals are same along all the three axes.

(vii) Hexagonal crystal : Two of the crystal axes are 60° apart while the third is perpendicular to both of them. The repetitive

intervals are the same along the axes that are 60° apart. However, the repetitive interval along the third axis is different.

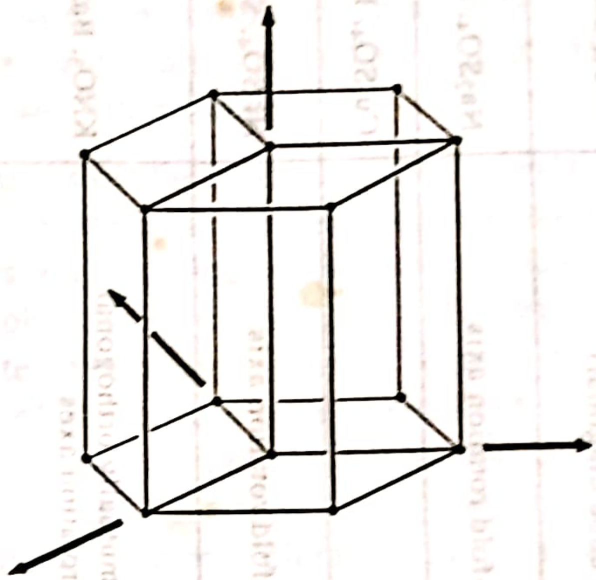


Fig. 19.5

The hexagonal system is so called because, a simple hexagonal structure can be formed from such unit cells as shown in Fig. 19.5. Instead of considering a hexagonal structure in terms of the rhombic cell of Fig. 19.4, it is more convenient to consider it in terms of a unit hexagonal cell like that of Fig. 19.5. A unit hexagonal cell has four axes, of which three, 120° apart, lie in one plane with the fourth axis perpendicular to this plane. The repetitive intervals along the three coplanar axes are equal, but is different along the fourth. From the discussion on the characteristic features of the different crystal systems, it is clear that the cubic system possesses the greatest symmetry whereas triclinic system has the least.

The characteristic features of the different crystal systems are summarized in table - 1.

Table - 1. The seven crystal systems, divided into 14 Bravais lattices

System	Bravais lattice	Unit cell characteristics		Characteristic symmetry elements	Examples
Cubic	Simple Body - centred Face - centred	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$	Four 3-fold rotation axes (along cube diagonal)	$NaCl$, CaF_2, $NaClO_2$
Monoclinic	Simple Base - centred	$a \neq b \neq c$	$\alpha = \beta = 90^\circ \neq \gamma$	One 2-fold rotation axis	Na_2SO_4, $FeSO_4$
Triclinic	Simple	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90^\circ$	None	$Cu\ SO_4$, $K_2Cr_2O_7$
Tetragonal	Simple Body - centred	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$	One 4-fold rotation axis	$Ni\ SO_4$, SnO_2
Orthorhombic	Simple Base - centred Body - centred Face - centred	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$	Three mutually orthogonal 2-fold rotation axes	KNO_3, $BaSO_4$, $MgSO_4$
Rhombohedral (Trigonal)	Simple	$a = b = c$	$\alpha = \beta = \gamma \neq 90^\circ$	One 3-fold rotation axis	$Ca\ SO_4$
Hexagonal	Simple	$a = b \neq c$	$\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$	One 3-fold rotation axis	SiO_2, AgI

19.5 The symmetry of crystals

What is meant by symmetry ? Formally, it may be said that a symmetry of a particular kind exists, when a certain operation leaves something unchanged. A candle is said to be symmetric about a vertical axis because it can be rotated about that axis without change in appearance or any other feature.

The very forms of the crystals exhibit a certain symmetry. This symmetry of the crystals is revealed from the measurement of their interfacial angles which do not vary from specimen to specimen, by observing the phenomenon that crystals tend to cleave along perfectly definite directions, from a study of the variation of refractive index of the crystal with direction inside the crystal and from the phenomenon of double refraction exhibited by certain optically active crystals. A study of crystals by such methods has shown that the main symmetry elements of a crystalline solid are (a) axes of symmetry (b) planes of symmetry and (c) centre of symmetry.

Consider a simple geometrical solid body such as a cube. Since there is a certain regularity in the arrangement of its faces, it may be called a symmetrical solid. In fact, the description of symmetries present in a solid is a way of describing the regularity with which the crystal faces are arranged.

When a cube is laid flat on a table, its four side faces become vertical, as shown in Fig. 19.6. The intersection of these four faces are

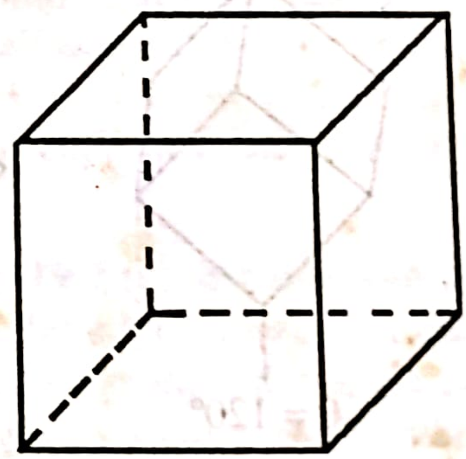


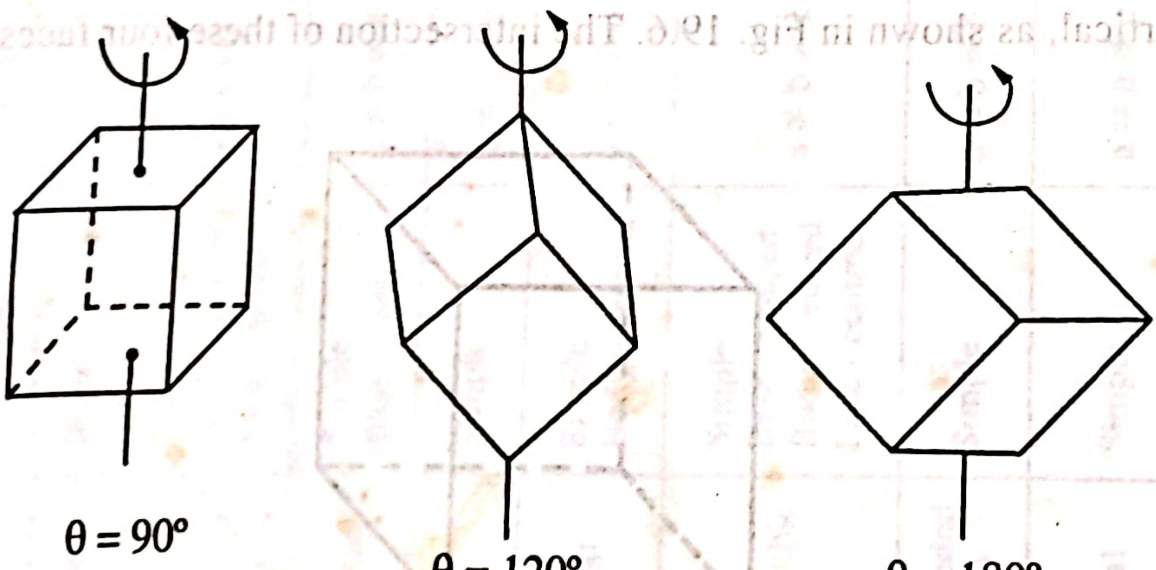
Fig. 19.6

four vertical lines which are parallel to each other. It is quite obvious that if one of the intersection edges is made vertical, then the four faces automatically become vertical planes. Thus the conclusion which can be drawn from this fact is that the faces are related to one another in some way. The four faces are said to constitute a *zone* and any direction parallel to the intersection edge is called the *zone axis*.

(a) axis of symmetry

Let an axis pass through the centre of the horizontal faces of the cube in a direction parallel to the intersection edges. Then, as shown in Fig. 19.7 (a), if the cube is rotated about this vertical axis, then in one complete revolution of 360°, there will be found four positions of the cube which are coincident with its original position. In fact, when the cube is rotated through 90°, it is not possible to distinguish it from its original position. In other words, each rotation of 90° brings the cube into self-coincidence or in a *congruent* position.

This rotation axis is known as an axis of symmetry. Since there are four congruent positions in one complete revolution, it is called a four-fold axis of symmetry. In general, if a rotation through an angle of $2\pi/n$ radians or $360/n$ degrees about an axis brings a figure into congruent position, the axis is called an n-fold axis of symmetry.



$\theta = 90°$

(a)

$\theta = 120°$

(b)

$\theta = 180°$

(c)

Fig. 19.7

If n = 1, the crystal has to be rotated through 360° to achieve self-coincidence. Such an axis is referred to as an *identity* axis. Every crystal has an infinite number of such axes.

If n = 2, the crystal has to be rotated through 180° for self-coincidence and the axis is referred to as a *diad* axis.

If n = 3, the crystal has to be rotated through 120° for self-coincidence and the axis is called a *triad*.

If n = 4, the angle of rotation has to be 90° and the axis is called a *tetrad*.

If n = 6, the angle of rotation has to be 60° and the axis is called a *hexad*.

It has been found that crystalline solids exhibit only 1,2,3,4 and 6-fold symmetry. They do not exhibit any 5-fold or any symmetry higher than 6. This is because, a crystal is not just a solid body, but a body in which the atoms and molecules are internally arranged in very regular and periodic fashion in a three-dimensional pattern. Consequently, identical repetition of a unit can take place only when we consider 1,2,3,4 and 6-fold axes.

In view of the above discussion, it can be seen that a cube possesses three axes of four-fold symmetry – one normal to each of the three pairs of parallel faces. In addition, it possesses the following axes of symmetry:

As shown in Fig. 19.7 (b), the cube can be rotated about a solid diagonal and a rotation through 120° about this diagonal brings it into a position of self-coincidence or congruence. Hence it is an axis of *3-fold symmetry i.e.*, a *triad*. Since there are, in all, *four solid diagonals in a cube, a cube has four axes of 3-fold symmetry i.e., four triads*.

As shown in Fig. 19.7 (c), the line joining the middle points of a pair of opposite parallel edges provides a *diad* axis, i.e., a rotation of 180° about this axis brings the cube into a position of congruence. Since there are 12 edges in a cube, the number of diad axes is 6. A cube, therefore, has 6 diad axes.

Thus, a cube has obviously the following 13 axes of symmetry:

3 tetrad axes, 4 triad axes and 6 diad axes.

(b) Planes of symmetry

Another type of symmetry possessed by a crystal is known as *plane of symmetry*. A crystal can be divided into two equal halves by an imaginary plane, such that one is the *mirror* image of the other *i.e.*, if one of the two halves is placed on a mirror, the image reproduces the other half of the crystal. Such a plane is called a plane of symmetry. As shown in Fig. 19.8, there are *three* planes of symmetry parallel to the faces of the cube.

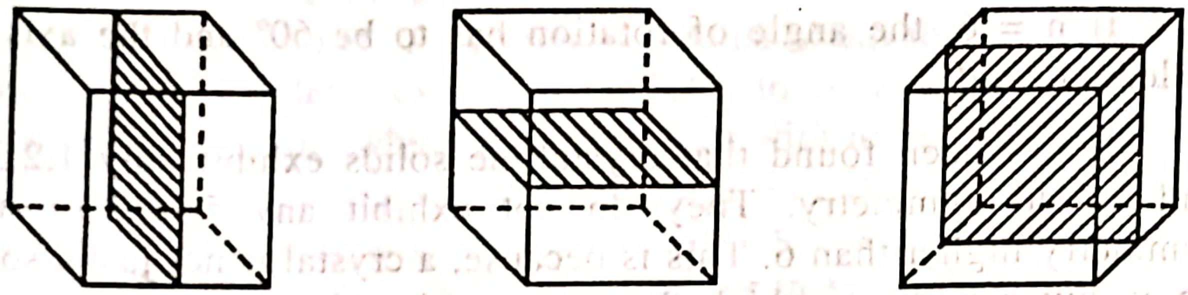
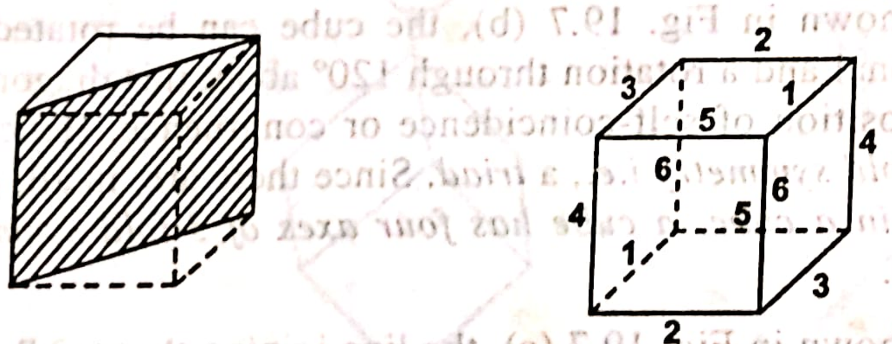


Fig. 19.8

In addition to the three planes of symmetry mentioned above, a cube also possesses 6 *diagonal* planes of symmetry one of which is shown in Fig. 19.9 (a). This plane is formed by a pair of opposite parallel edges. As there are 6 such pairs of edges [shown by the same



(a) (b)

Fig. 19.9

numbers in Fig. 19.9 (b), the number of diagonal planes of symmetry is six. Thus the total number of planes of symmetry that can be possessed by a cubic crystal is (3+6) = 9.

(c) Centre of symmetry

There is yet another type of symmetry, called the *centre of symmetry*, that a crystal possesses. It is a point such that if a line is drawn from any point on the crystal through this point and produced an equal distance on the other side of this centre, it meets an identical point. Hence, it is also known as the *centre of inversion* (I). It is important to mention here that a crystal may possess a number of planes or axes of symmetry but it can have only one centre of symmetry.

To sum up, the full crystallographic symmetry of a cubic crystal comprises of the following 23 elements of symmetry.

(i) 13 axes of symmetry: 3 tetrads, 4 triads and 6 diads.

(ii) 9 planes of symmetry: 3 planes parallel to the faces and 6 diagonal planes.

(iii) 1 centre of symmetry.

The unit cells belonging to the different crystal systems have the following symmetry axes.

(i) the triclinic system has no axis of symmetry.

(ii) the monoclinic system has one diad (2-fold rotation) axis only.

(iii) the orthorhombic system has three mutually orthogonal diad (2-fold rotation) axes only.

(iv) the rhombohedral (trigonal) system has one 4-fold rotation axis.

(v) the cubic system has four triad (3-fold rotation) axes.

(vi) the tetragonal system one triad (3-fold rotation) axis which is equally inclined to the three co-ordinate axes.

(vii) the hexagonal system has one hexad which is at right angles to the other co-ordinate axes.

19.6 Space lattices of Cubic System

A cubic structure is the simplest type of array in which the atoms take positions at the corners of the cube. However, depending

on the positions of the lattice points in the unit cell, three different types of lattices are possible in this system. These are

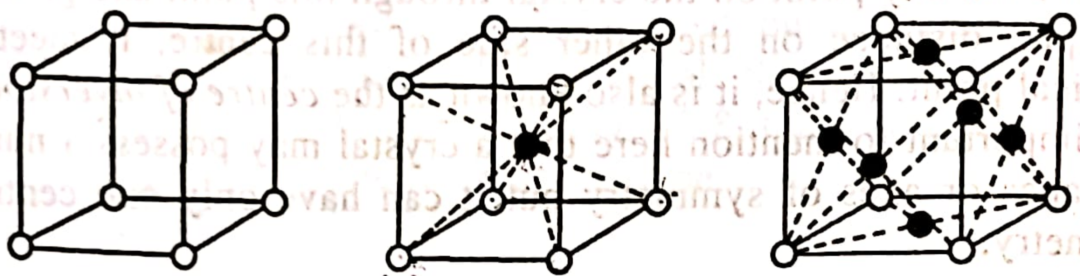


Fig. 19.10

(i) Simple cubic (SC) lattice : this has points only at the corners. [Fig. 19.10 (a)]

(ii) Body-centred cubic (BCC) lattice : in addition to points at the corners, this has one additional point at the centre of the cell [Fig. 19.10(b)].

(iii) Face-centred cubic (FCC) lattice : in addition to pints at the corners, a face-centred cubic lattice has six additional points, one on each face [Fig. 19.10(c)].

The SC, BCC and FCC lattices are sometimes represented by the symbols P,I and F respectively. Not all such lattices are possible in all crystal systems. For example, a cubic system has all three lattices *i.e.*, P,I and F whereas triclinic and hexagonal systems have P lattice only.

19.7 Co-ordination number

There is an interesting point about cubic lattices. This is called the *co-ordination number. It is defined as the number of nearest neighbours which an atom has in the unit cell of any crystal structure.* It is characteristic of a given space lattice and is determined by an inspection of the model. The following cases would be considered :

(i) SC structure : In this case, each corner atom is linked with seven other unit cells that can be imagined to be built around the unit cell containing the atom. In that case, each

corner atom has four neighbours in the same plane, one vertically above and one immediately below, giving a total of six *nearest* neighbouring atoms. Hence, in the case of SC structure the co-ordination number is *six.*

(ii) BCC structure : In the case of BCC structure, the nearest neighbours of any corner atom are the body-centred atoms and not the other corner atoms. Since, there are eight surrounding unit cells for any corner atom, their eight body-centred atoms form the nearest neighbours for any corner atom. Hence the co-ordination number of a BCC structure is *eight.*

(iii) FCC structure : In the case of FCC structure the *nearest* neighbours of any corner atom are the face-centred atoms of the surrounding unit cells. Any corner atom has four such atoms in its own plane, four in a plane above it and four in a plane below it. Hence the co-ordination number of such a structure is *twelve.*

19.8 Number of atoms per unit cell 2012

(i) Simple cube

Each corner atom in a simple cube is shared by *eight* surrounding cubes. The share of each cube, therefore, comes to one-eighth of an atom [Fig. 19.11(a)]. As there are eight corner atoms in all, each cube has $8 \times {}^{1}/_{8} = 1$ atom.

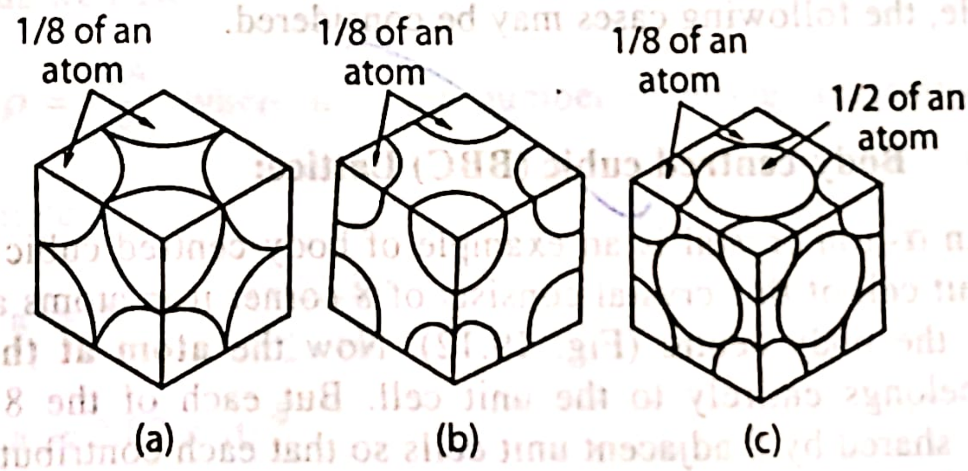


Fig. 19.11

(ii) Body-centred cube

A body-centred cube has eight corner atoms and one centre atom [Fig. 19.11(b)]. Thus the number of atoms which belong to this cube are

(a) 1 centre atom

(b) $8 \times {}^1/_8 = 1$ corner atom.

Hence, number of atoms per cube = 1 + 1 = 2.

(iii) Face-centred cube

A face-centred cube has eight corner and six face-centred atoms [Fig. 19.11(c)]. Each of the six face-centred atoms is shared by the two adjoining cubes. Thus the number of atoms which belong to this cube are

(i) 6/2 = 3 face-centred atom

(ii) $8 \times {}^1/_8 = 1$ corner atom.

Hence, total number of atoms per cube = 3 + 1 = 4.

19.9 Calculation of dimensions of unit cell

The dimensions of the unit cell or the interatomic distance in a crystal lattice can be computed from a knowledge of (i) its crystalline form, (ii) density of the material, (iii) molecular weight of the crystalline compound and (iv) Avogadro's number. As an example, the following cases may be considered.

(i) Body-centred cubic (BBC) Lattice:

An α-iron crystal is an example of body-centred cubic lattice. Each unit cell of this crystal consists of 8 corner iron atoms and one atom at the body centre (Fig. 19.12). Now the **atom at the body** centre belongs entirely to the unit cell. But each of the 8 corner atoms is shared by 8 adjacent unit cells so that each contributes one-eighth of an atom to the cell under consideration. Hence, total number of iron atoms associated with each unit cell = 1 + (8/8) = 2.

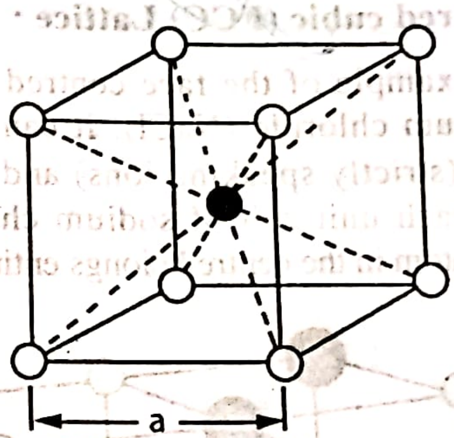


Fig. 19.12

Atomic weight of iron = 55.85 and its density, ρ, is 7.86 gm/cm^3.

If a cm is the length of the edge of the unit cell, then its volume is a^3 and the mass of unit cell is $a^3\rho = a^3 \times 7.86$.

Again, since the atomic weight is 55.85, the mass of each atom

$= \frac{A}{N} = \frac{55.86}{6.02\times10^{23}}$ where $N = 6.02 \times 10^{23}$ is the Avogadro's number

and A the atomic weight.

Hence, the mass of 2 atoms.

$$= \frac{2\times55.86}{6.02\times10^{23}}$$

Thus we have

$a^3\rho = \frac{nA}{N}$ where n is the number of atoms associated with each unit cell.

$$\text{or, } a^3 \times 7.86 = \frac{2\times55.86}{6.02\times10^{23}}$$

$$\text{or, } a = 2.87 \times 10^{-8}$$

$$= 2.87\ A^{\circ}$$

(ii) Face-centred cubic (FCC) Lattice :

The common example of the face-centred cubic lattice system is the crystal of sodium chloride (NaCl). It can be proved that four sodium (Na^+) atoms (strictly speaking ions) and four chlorine atoms are associated with each unit cell of sodium chloride. As shown in Fig. 19.13, the sodium atom in the centre belongs entirely to the cell shown.

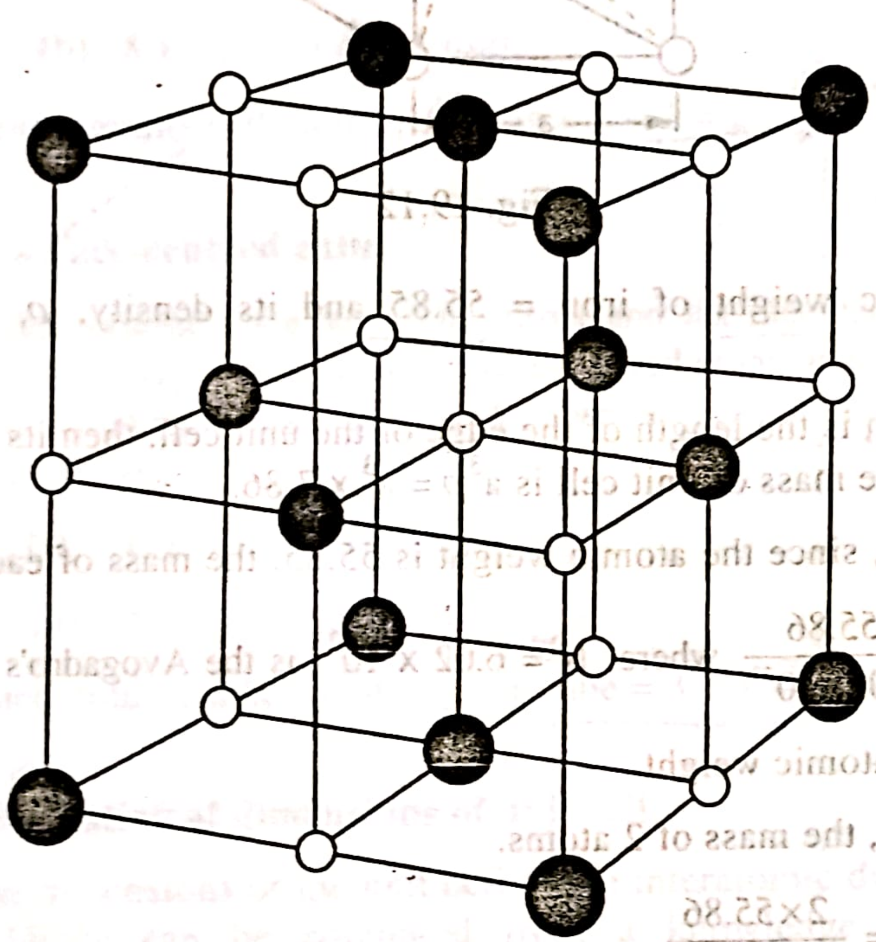


Fig. 19.13

Each of the other twelve sodium atoms is shared with three adjacent unit cells so that each contributes one-fourth of an atom to the cell under consideration. Hence, total number of Na atoms in each unit cell is

$$= 1 + \frac{1}{4}(12) = 4$$

Now, as for the chlorine atoms, there are (i) 8 corner atoms and (ii) 6 face atoms - one in each face. It can be easily visualized that each one of the corner chlorine atoms is shared by 8 adjacent unit

cells so that each contributes one-eighth of an atom to the cell under consideration. Hence, number of corner atoms per unit cell is = 8/8 = 1.

Considering the face-centred chlorine atoms, each one of these atoms is shared by two unit cells. Hence, each contributes one-half of an atom to the cell under consideration.

Thus, the number of face-centred chlorine atoms associated with each unit cell = 6/2 = 3. Hence, the total number of chlorine atoms belonging to each unit cell of NaCl is = 1 + 3 = 4. Taken together, it means that there are 4 *molecules* of NaCl per unit cell.

Molecular weight of NaCl = 23 + 35.5 = 58.5

Density of NaCl = 2.18 gm/cm^3

If a is the length of the unit cell, ρ the density of NaCl, n the number of molecules per unit cell, then we have as in the case of FCC lattice,

$$a^3\rho = n\ M/N$$

where M is the molecular weight and N the Avogadro's number.

$$a^3 \times 2.18 = \frac{4\times 58.5}{6.02\times 10^{23}}$$

or, $a = 5.63 \times 10^{-8}$ cm = 5.63 A°.

As can be seen from Fig. 19.13, a is the distance between two adjacent atoms of the same kind – chlorine atoms in this case. The distance between two adjacent atoms, *i.e.*, Na and Cl atoms is obviously half of this value, *i.e.*, d = 5.63/2 = 2.815 A°.

19.10 Atomic radius

It is assumed that the atoms in an elemental solid are spherical in nature and that any two neighbouring atoms touch each other. The atomic radius is defined as half the distance between two nearest neighbouring atoms in a crystal of pure element. It is easy to calculate the atomic radii from a knowledge of the crystal structure of the solid and the lattice parameters. Usually the atomic radius is expressed in terms of the cube edge a.

(i) s.c. structure

In Fig. 19.14, the edge of the cube is a. By definition, the atomic

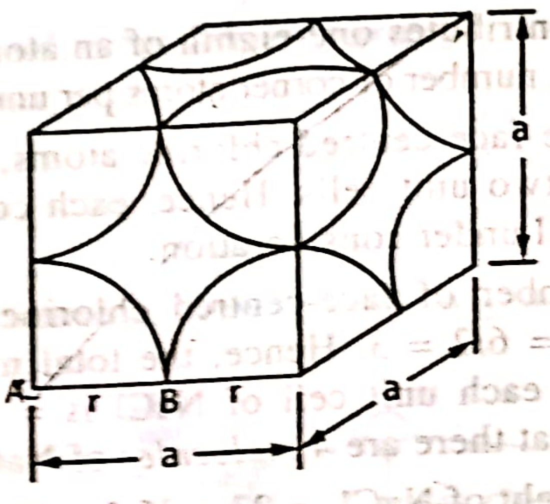


Fig. 19.14

radius is AB = r. It is obvious from the figure,

$$r = a/2.$$

(ii) b c c structure

From Fig. 19.15

$$AB^2 = a^2 + a^2 = 2a^2$$

$$AC^2 = AB^2 + BC^2$$

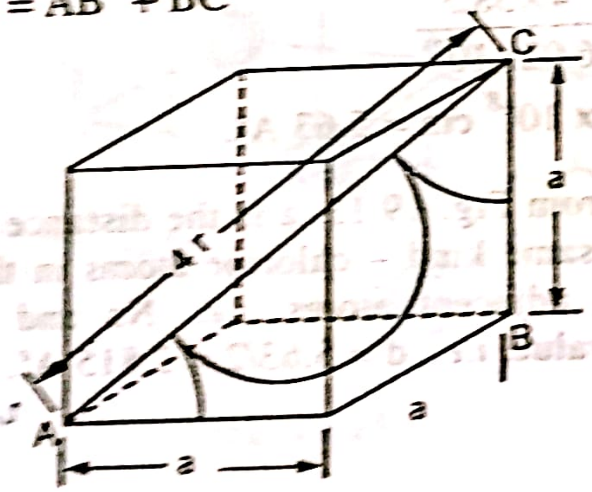


Fig. 19.15

$$= (2a)^2 + a^2 = 3a^2$$

$$\therefore \; AC = \sqrt{3}\,.\,a$$

As shown in the diagram AC = 4r.

$$\therefore \; 4r = \sqrt{3}\,.\,a$$

$$\text{or, } r = \frac{\sqrt{3}}{4}\,.\,a$$

(iii) fcc structure

As seen from Fig. 19.16, the surface diagonal AC equals four atomic radii.

Now $AC^2 = AB^2 + BC^2 = a^2 + a^2 = 2a^2$

or, $AC = \sqrt{2} \cdot a$

Also $AC = 4r$.

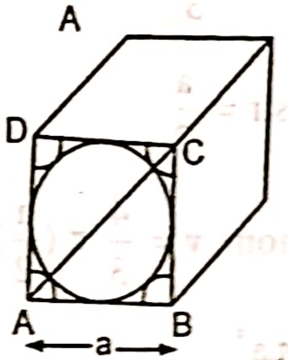


Fig. 19.16

$\therefore\ 4r = \sqrt{2} \cdot a$

or, $r = \frac{\sqrt{2}}{4} \cdot a = \frac{1}{2\sqrt{2}} \cdot a$

The lattice constant or side of a cubic unit cell as well its face and body diagonals in terms of atomic radius are listed in Table. II

TABLE II

	Side	Face diagonal	Body diagonal
S. C.	$a = 2r$	$\frac{1}{\sqrt{2}}(4r)$	$\frac{\sqrt{3}}{2}(4r)$
B. C. C.	$a = \frac{4}{\sqrt{3}} \cdot r$	$\sqrt{\frac{2}{3}}(4r)$	$4r$
F. C. C.	$a = 2\sqrt{2} \cdot r$	$4r$	$\sqrt{\frac{3}{2}}(4r)$

19.11 Packing fraction

The relative density of packing or the packing fraction is defined as the ratio of the actual volume occupied by the spherical atoms to the total available volume of the structure.

(i) sc structure

As shown earlier, this cube has one atom.

$$\text{Volume of one atom} = \frac{4}{3}\pi r^3$$

$$\text{Now, atomic radius } r = \frac{a}{2}$$

$$\therefore \text{ volume of the atom, } v = \frac{4}{3}\pi\left(\frac{a}{2}\right)^3$$

$$= \frac{\pi . a^3}{6}$$

$$\text{volume of the cube, } V = a^3$$

$$\therefore \text{ packing fraction, } p = \frac{v}{V} = \frac{\pi . a^3}{6a^3} = \frac{\pi}{6}$$

(ii) b c c structure

$$\text{volume of one atom} = \frac{4}{3}\pi r^3$$

$$\text{but } r = \left(\frac{\sqrt{3}}{4}\right) . a$$

$$\therefore \text{ volume of one atom} = \frac{4}{3}\pi\left[\frac{\sqrt{3}}{4}\right]^3 . a^3$$

For bcc structure, the number of atoms per cube = 2.

$$\therefore \quad v = 2 \times \frac{4}{3}\pi\left[\frac{\sqrt{3}}{4}\right]^3 . a^3$$

$$= \sqrt{3}\left[\frac{\pi}{8}\right] \cdot a^3$$

volume of the cube $\qquad V = a^3$

$$\therefore \text{ packing fraction, } p = \frac{v}{V} = \frac{\sqrt{3}\left[\frac{\pi}{8}\right].a^3}{a^3}$$

$$= \frac{\sqrt{3}\pi}{8}$$

(iii) fcc structure

$$\text{atomic radius, } r = \frac{a}{2\sqrt{2}} = \frac{\sqrt{2}.a}{4}$$

$$\therefore \text{ volume of one atom} = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi\left[\frac{\sqrt{2}a}{4}\right]^3$$

For fcc structure, the number of atoms per cube = 4.

$$\therefore\ v = 4 \times \frac{4}{3}\pi\left[\frac{\sqrt{2}a}{4}\right]^3 = \sqrt{2}\pi\frac{a^3}{6}$$

volume of the cube, $V = a^3$

$$\therefore \text{ packing fraction, } p = \frac{v}{V}$$

$$= \frac{\sqrt{2}\pi a^3}{6a^3} = \frac{\sqrt{2}}{6}\pi.$$

$$= \sqrt{2}\left(\frac{\pi}{6}\right)$$

Sum of the details of elementary cubic structure are summarised in the table below.

Crystal Structure	Co-ordination number	Atomic radius	Atoms per unit cell	Packing fraction
SC	6	$r = \frac{a}{2}$	1	$\pi/6$
BCC	8	$r = \sqrt{3}\left(\frac{a}{4}\right)$	2	$\sqrt{3}\left(\frac{\pi}{8}\right)$
FCC	12	$r = \sqrt{2}\left(\frac{a}{4}\right)$	4	$\sqrt{2}\left(\frac{\pi}{6}\right)$

19.12 Important plane systems in a cubic crystal

In a cubic crystal, three sets of planes are very rich in atoms. Consequently, Bragg or X-ray reflections from these planes are more intense than from others. These three sets of planes, having different spacings are shown in Fig. 19.17.

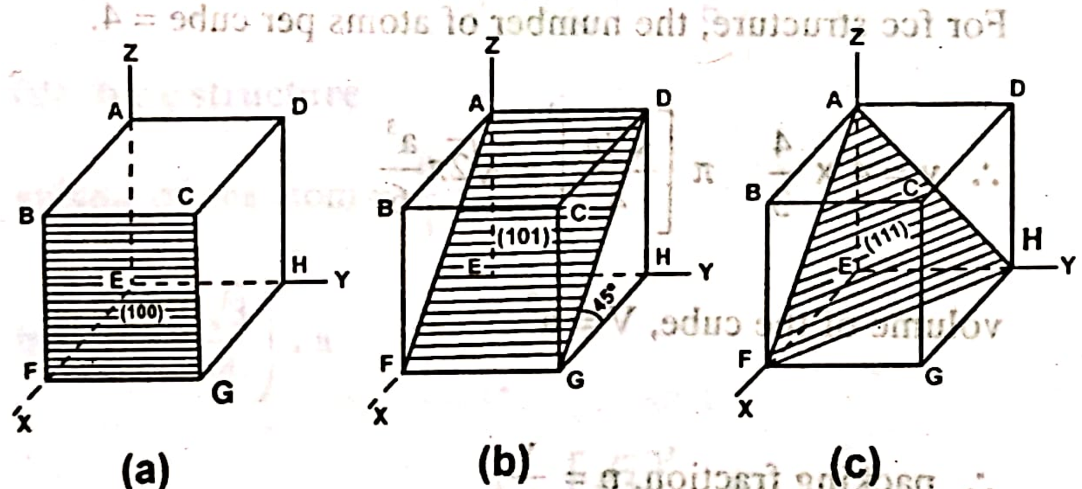


Fig. 19.17

(i) The first set of planes consists of the surface planes of the cube and those parallel to them. Planes BFGC, AEHD, ABCD, EFGH, etc. in Fig. 19.17 (a) represent these planes. Let the spacing between consecutive parallel surface planes be d_1 as shown in Fig. 19.18.

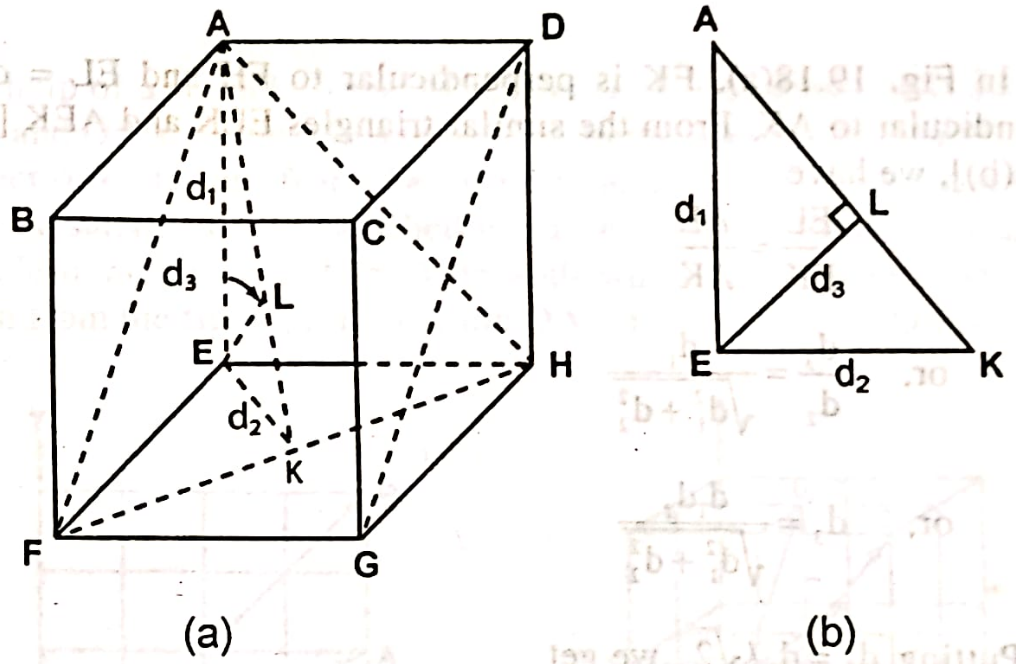


Fig. 19.18

(ii) The second set of planes are parallel planes which are inclined at an angle of 45° to the planes mentioned above. Planes like AFGD, etc. in Fig. 19.17 (b) represent these planes. Let the spacing between two consecutive planes of this class be d_2 as shown in Fig. 19.18.

(iii) Planes which are parallel to the plane AFH [Fig. 19.17 (c)] represent the third set of planes. Let the spacing between these planes be d_3 as shown in Fig. 19.18.

Spacing of planes :

For each basic crystal structure, a definite ratio exists between the spacings of these planes. It is, therefore, essential to have a knowledge of these spacings in order to identify different crystal structures.

Fig. 19.17 shows the case of a simple cubic crystal with atoms at the corner of the cubes. The spacing between planes like ABCD and EFGH is d_1. The spacing of planes like AFGH which are inclined at 45° to the planes like ABCD is $d_2 = d_1 / \sqrt{2}$. The spacing d_3 between the third set of planes like AFH is found by drawing the triangle AEK in true shape as in Fig. 19.18 (b).

In Fig. 19.18(a), EK is perpendicular to FH and EL = d_3 is perpendicular to AK. From the similar triangles ELK and AEK [Fig. 19.18(b)], we have

$$\frac{EL}{EK} = \frac{AE}{AK}$$

or, $$\frac{d_3}{d_2} = \frac{d_1}{\sqrt{d_1^2 + d_2^2}}$$

or, $$d_3 = \frac{d_1 d_2}{\sqrt{d_1^2 + d_2^2}}$$

Putting $d_2 = d_1/\sqrt{2}$, we get

$$d_3 = \frac{d_1^2}{\sqrt{2}\sqrt{d_1^2 + d_1^2/2}} = \frac{d_1^2}{\sqrt{2}\sqrt{\frac{2d_1^2 + d_1^2}{2}}} = \frac{d_1}{\sqrt{3}}.$$

$$\therefore \frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} : : 1 : \sqrt{2} : \sqrt{3}$$

Similarly it can be shown that for a body-centred lattice (bcc) structure

$$\frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} : : 1 : \frac{1}{\sqrt{2}} : \sqrt{3}$$

and for a face-centred lattice (fcc) structure,

$$\frac{1}{d_1} : \frac{1}{d_2} : \frac{1}{d_3} : : 1 : \sqrt{2} : \frac{\sqrt{3}}{2}$$

Bragg performed experiments on reflection of monochromatic X-rays from different crystals and proved the above ratios. He also used these ratios to determine whether the crystal was simple, cubic or body-centred type.

Direction, Planes and Axial units

The ability to describe relative orientations of lines and planes is a basic necessity for crystal analysis. This can be achieved with

the help of a suitable frame of reference. For example, two directions OA and OB are shown in a two-dimensional plane in Fig. 19.19. These directions, starting from the origin O and ending at points A and B, can be satisfactorily described by giving the co-ordinates of the first whole-numbered point (x,y) through which each passes. As can be seen from the figure, for direction OA it is (2,1) and for OB it is (1,1).

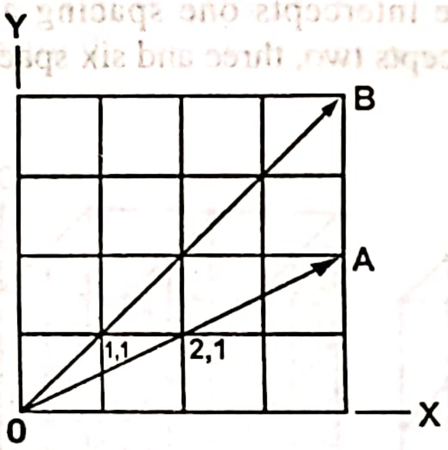


Fig. 19.19

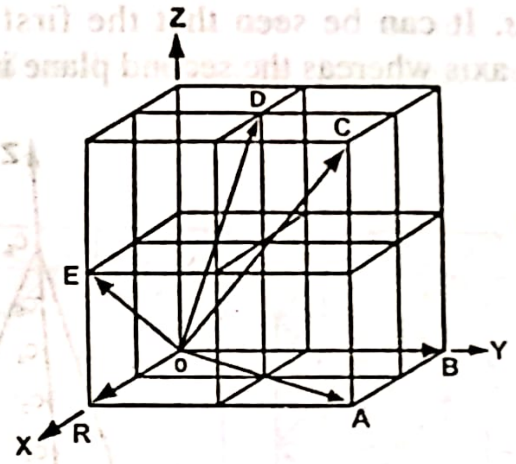


Fig. 19.20

Similarly, a direction in space, *i.e.*, in a three-dimensional plane can be described by giving the co-ordinates of the first whole-numbered point (x,y,z) through which they pass. The directions of the lines in Fig. 19.20 are given as follows

OA [110] OR [100]

OB [010] OD $[\frac{1}{2},\frac{1}{2},1]$ or [1 1 2]

OC [111] OE $[1,0,\frac{1}{2}]$ or [2 0 1]

The *square* brackets are used to indicate a *direction*. Commas are unnecessary but the digits are read as *one-zero-zero* or *one-aught-aught* and not *one hundred.* The digits in square brackets indicate the indices of that direction and are generally written as [*uvw*]. A bar over the digit is used to indicate a negative index. For example +X-axis has indices of [100] whereas – X- axis has $[\bar{1}00]$ indices. Similarly a +Y-axis is indicated by [010] while a – Y-axis is

indicated by $[0\bar{1}0]$. A +Z-axis is indicated by [001] while a - Z-axis is indicated by $[00\bar{1}]$.

Let the faces $a_1b_1c_1$ and $a_2b_3c_6$ make certain intercepts or segments on the reference axes OX, OY and OZ (Fig. 19.21). The intercepts made by the first plane are Oa_1, Ob_1 and Oc_1 while those made by the second plane are Oa_2, Ob_3 and Oc_6. These intercepts or segments represent the respective *linear parameters* of the two *faces*. It can be seen that the first face intercepts one spacing along each axis whereas the second plane intercepts two, three and six spacings

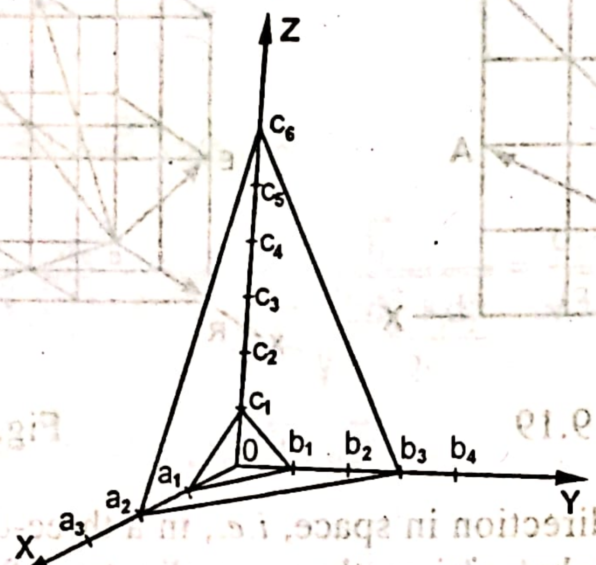


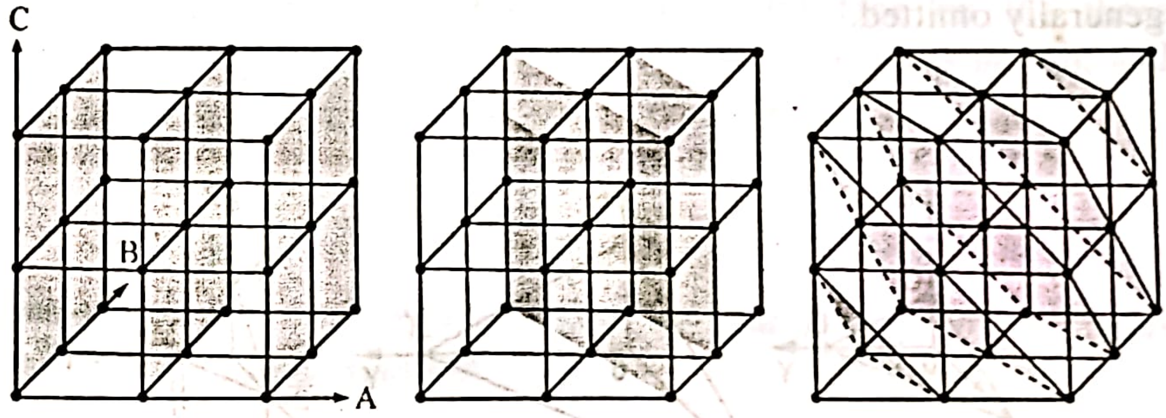
Fig. 19.21

along the respective axes. If the spacings of the first plane are 1a, 1b and 1c along the three axes, then the spacings made by the second plane along the three axes are 2a, 3b and 6c respectively. If the spacings of the first plane are related to each other as 1 : 1 : 1, then the spacings of the second plane are related to each other as 2 : 3 : 6. These figures of 1, 2, 3, 6 represent the *numerical parameters* of the *faces* and must always be whole numbers.

The first face is referred to as the *unit face* and may be defined *as the face which makes one or unit intercept on each axis*. Linear parameters intercepted by the unit *face* are called *axial units* and are generally represented by a, b and c along X, Y and Z axes respectively.

19.13 Miller Indices

Orientations of planes or faces in a crystal may also be described by a system that was first used in 1839 by W.H. Miller, an English mineralogist, to describe the faces of a crystal. As discussed earlier, a face or an internal plane of a crystal will have intercepts on the crystallographic axes *i.e.*, reference axes. All the series of planes which are parallel to this plane will have intercepts whose ratios to the axial lengths of the *unit* cell, a, b, and c are independent of the particular axial lengths involved in the given lattice. Since we are



(a) (100) planes (b) (110) planes (c) (111) planes

Fig. 19.22

concerned with the orientation of the plane and not its absolute position, these ratios could serve to specify the plane. But these ratios cause difficulty when the plane is parallel to a crystallographic axis as in Fig. 19.22. In this case the plane does not make any intercept on two of the axes at all. When the plane does not intersect any axis, then the intercept on this axis by the plane is said to be infinity. To avoid the introduction of infinity in order to specify the orientation of a plane, the reciprocal of the fractional intercept is used. Thus the reciprocal is zero when the intercept is infinity *i.e.*, when the plane and the axis are parallel. This leads to a system of determining the orientation of a plane in a crystal lattice by the *Miller indices*, which are defined as *a set of smallest integers in the ratio of the reciprocals of the fractional intercepts which the plane makes with the crystallographic axis*. For example, let us

consider the case of plane ABC of Fig. 19.23(a) which has intercepts of 2 axial units along OX, 2 axial units along OY and one axial unit along OZ. Hence the ratios of the axial lengths or the numerical parameters of the plane are 2, 2, and 1. Hence its orientation is given by (2,2,1). However, as suggested by Miller, it is more convenient to specify the orientation of the plane by the ratios of the reciprocal of the numerical parameters. The ratios of the reciprocals of the numerical parameters of the plane ABC are $\left(\frac{1}{2}:\frac{1}{2}:1\right)$. Hence, Miller indices of the planes are (1 1 2), the ratio sign (:) being generally omitted.

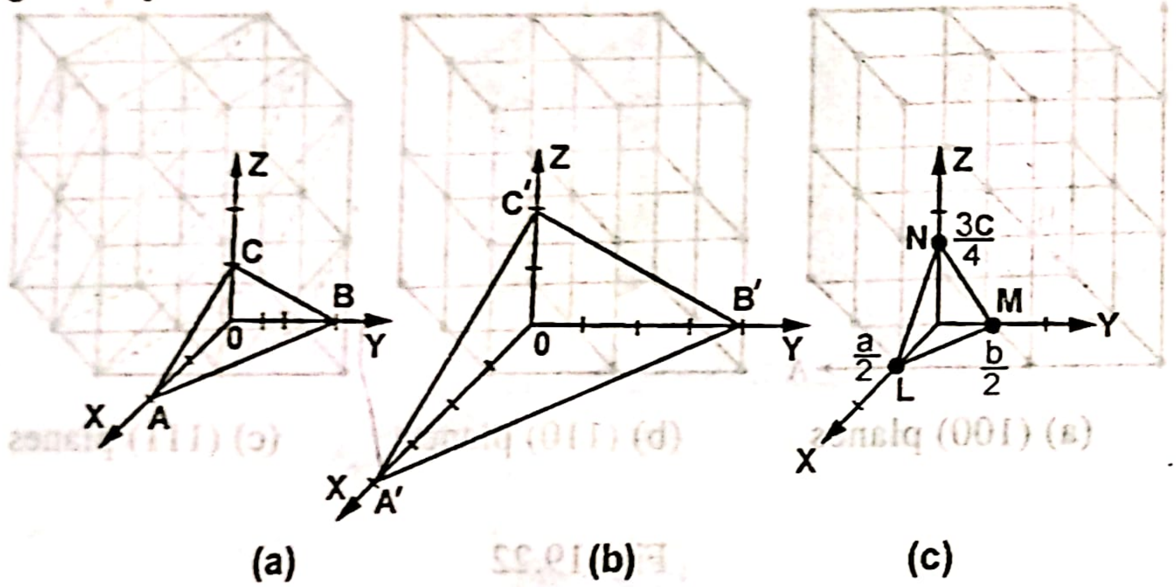


Fig. 19.23

Similarly, the numerical parameters of the plane A'B'C' [Fig. 19.23(b)] are 4, 4 and 2, the ratios of their reciprocals being $\left(\frac{1}{4}:\frac{1}{4}:\frac{1}{2}\right)$. Hence the Miller indices of the plane is (1 1 2) which shows that both ABC and A'B'C' belong to the same family of parallel planes.

Let us now consider the plane LMN [Fig. 19.23(c)].

The intercepts made by LMN when expressed in terms of axial units are a/2, b/2, 3c/4. Its numerical parameters are then

$\left(\frac{1}{2}, \frac{1}{2}, \frac{3}{4}\right)$. The ratios of the reciprocal of these numerical parameters are then $\left(\frac{1}{1/2} : \frac{1}{1/2} : \frac{1}{3/4}\right)$ or $\left(2 : 2 : \frac{4}{3}\right)$ or (6 6 4) or (3 3 2). Hence the Miller indices of this plane given by the smallest integers which are inversely proportional to the numerical parameters of the plane is (3 3 2).

It is customary to designate the Miller indices of a plane as (h k l) which means that the plane has fractional intercepts of 1/h, 1/k, and 1/l with the axes and that the actual intercepts are a/h, b/k and c/l. In general, if the intercepts made by a crystal face on the three reference axes are expressed in terms of the axial units pa, qb, and rc, then the Miller indices of the face are given by

$$\left(h : k : l :: \frac{1}{p} : \frac{1}{q} : \frac{1}{r}\right)$$

with the condition that h, k and l are the smallest possible integers satisfying the above equation.

Miller Indices of Cubic Crystal Planes

While determining the Miller indices of a plane, the following points should be kept in mind:

(i) When a plane is parallel to one of the reference axis, it is said to intersect that axis at infinity. Hence its Miller index for that axis is zero, since $1/\infty = 0$.

(ii) When a plane makes an intercept on the negative side of the axis, the Miller index for the axis is expressed by putting a negative sign directly over the index.

For examples, the shaded planes of Fig. 19.17 may be considered. The face plane BFGC [Fig. 19.17(a)] has an intercept on the X-axis but is parallel to both Y-and Z-axes. Hence it has Miller indices of (100). Similarly the plane AFGD [Fig. 19.17(b)] has intercepts on X-and Z-axes but not on Y-axis. Hence it has Miller indices of (101). The plane AFH [Fig. 19.17(c)] has equal intercepts on all the three axes. Hence it has a Miller indices of (111).

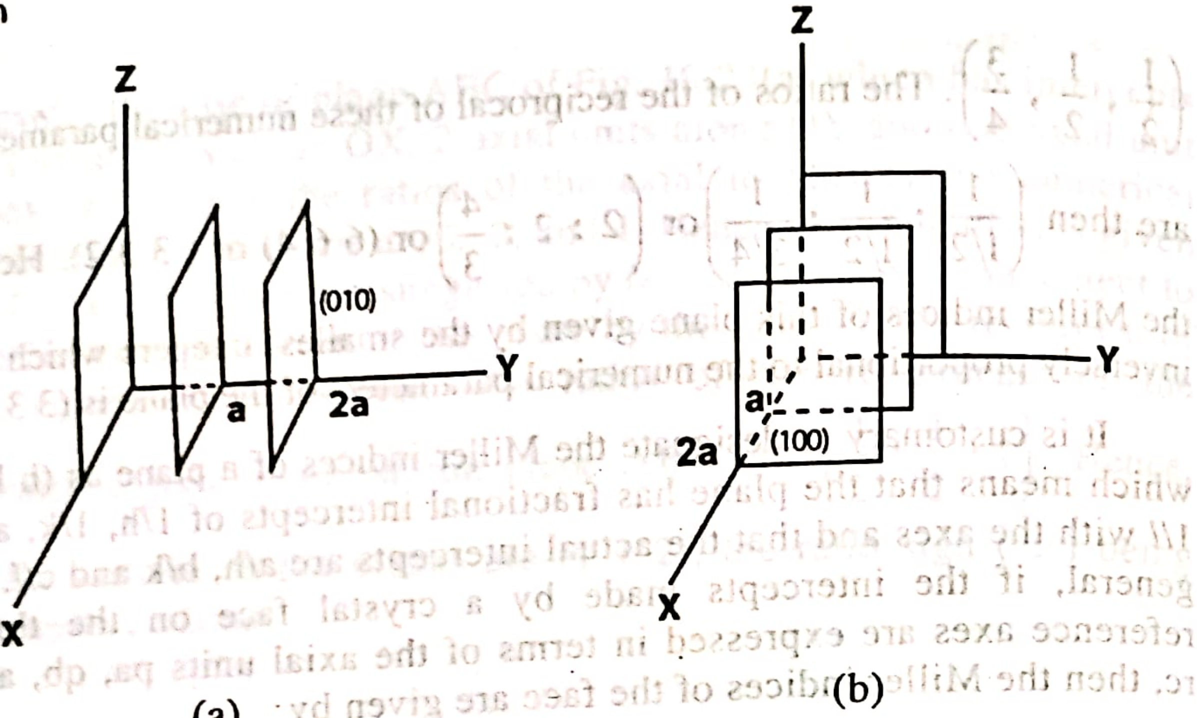


Fig. 19.24

Planes parallel to each other have the same Miller indices.

Fig. 19.24(b) shows a set of parallel planes normal to the X-axis. The Miller indices of these planes are designated (100). Similarly, a set of parallel planes normal to the Y-axis are sketched in Fig. 19.24(a). These planes have Miller indices of (010). Fig. 19.25(a) shows the sketch

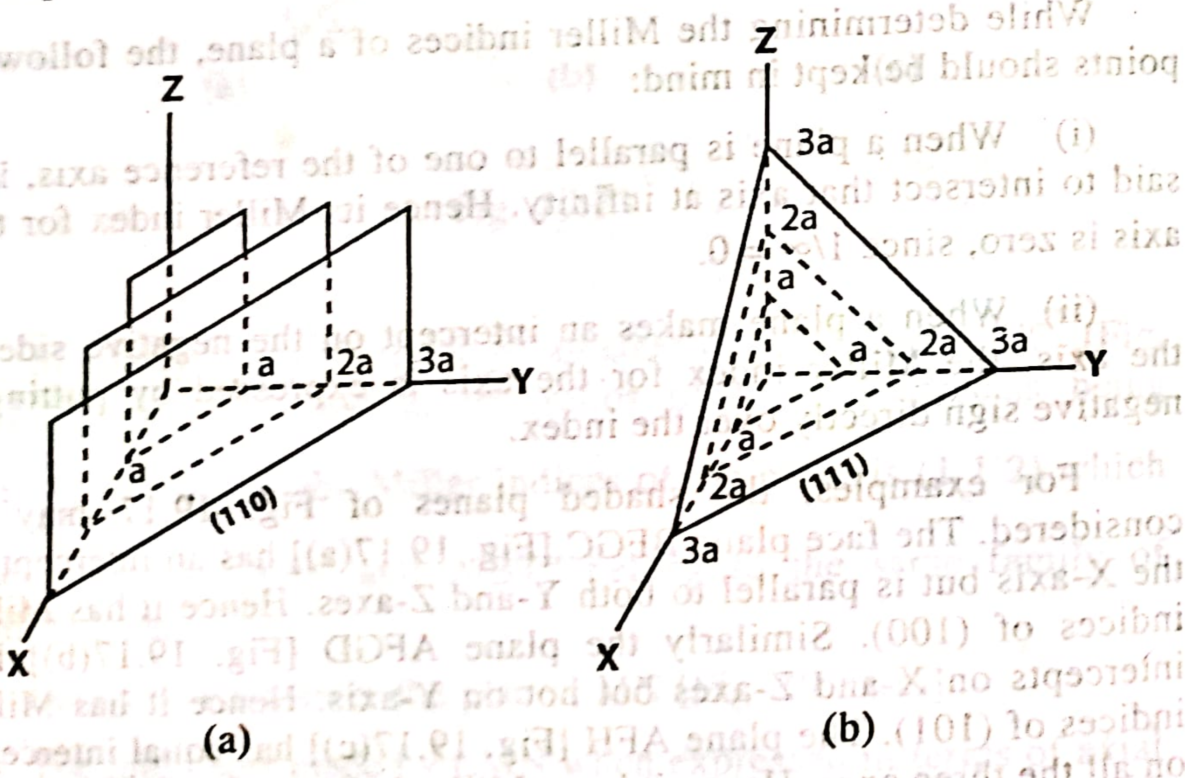


Fig. 19.25

of a set of parallel planes having equal intercepts on X-and Y-axes but parallel to the Z-axis. The Miller indices of these planes are designated (110). The planes shown in Fig. 19.25(b) have equal intercepts on all the three axes. Hence, the Miller indices of all these planes are (111). The family of planes shown in Fig. 19.26(b) have intercepts on the negative side of the X-axis and equal but positive

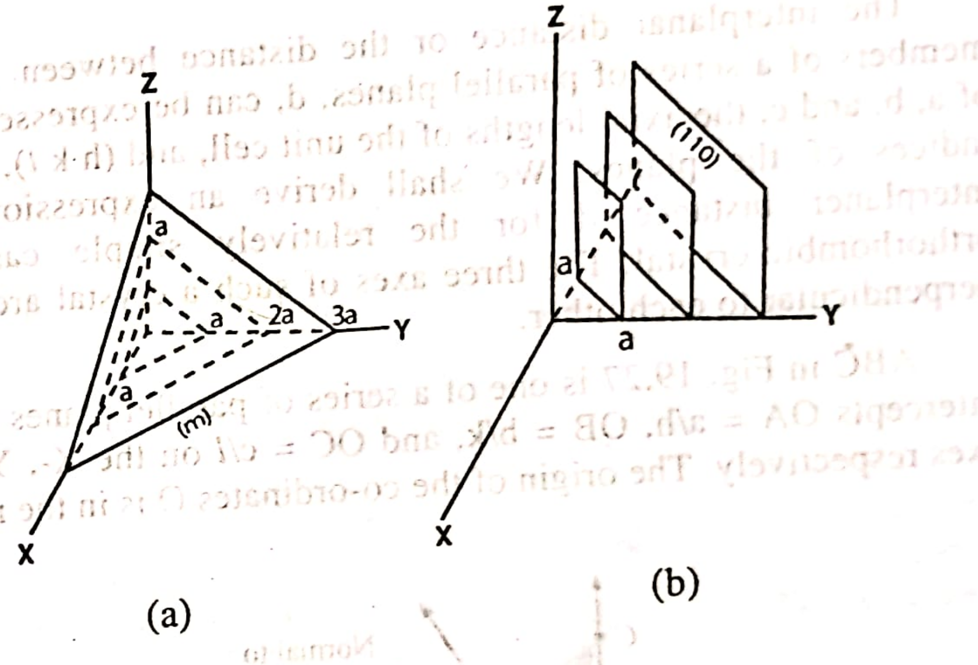


Fig. 19.26

intercept on the Y-axis. Hence their Miller indices are designated as $(\bar{1}\,10)$.

Procedure for Finding Miller Indices

(i) First, find the intercepts made by the plane on the X- , Y- and Z-axes.

(ii) Next, express them in terms of the axial units. Let these be x = a, y = 2b and z = 3c (or in general terms, x = pa, y = qb and z = rc).

(iii) Then, take the numerical parameters of the plane *i.e.*, 1 : 2 : 3 (p : q : r) and find the raio of their reciprocals $\left(\frac{1}{1} : \frac{1}{2} : \frac{1}{3}\right)$ or

$$\left(\frac{1}{p} : \frac{1}{q} : \frac{1}{r}\right).$$

(iv) Now, convert these reciprocals into smallest possible integers by multiplying each with their LCM.

In the present case the LCM = 6. Hence the reciprocals when converted to whole numbers become (6 : 3 : 2). Omitting the : sign, the Miller indices of the plane is designated as (632).

Interplanar distance

The interplanar distance or the distance between successive members of a series of parallel planes, d, can be expressed in terms of a, b, and c, the axial lengths of the unit cell, and (h k l), the Miller indices of the planes. We shall derive an expression for the interplaner distance d for the relatively simple case of an orthorhombic crystal. The three axes of such a crystal are mutually perpendicular to each other.

ABC in Fig. 19.27 is one of a series of parallel planes which has intercepts OA = a/h, OB = b/k, and OC = c/l on the X-, Y-, and Z-axes respectively. The origin of the co-ordinates O is in the next plane

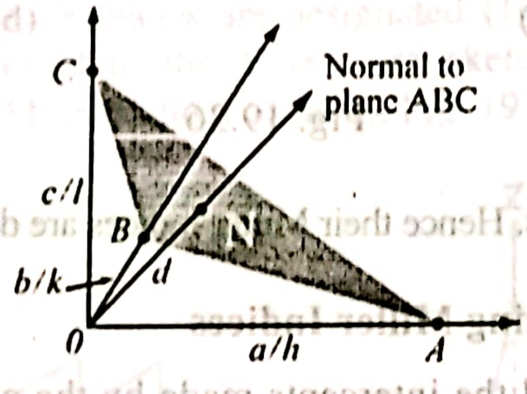


Fig. 19.27

of the set parallel to ABC. Therefore, ON, the length of the normal from the origin to the plane, is equal to d. Let α, β and γ be the angles that ON makes with the three crystallographic axes respectively. Then the direction cosines of ON are

$$\cos\alpha = \frac{ON}{OA} = \frac{d}{a/h}$$

$$\cos\beta = \frac{ON}{OB} = \frac{d}{b/k}$$

$$\cos\gamma = \frac{ON}{OC} = \frac{d}{c/l}$$

Since the sum of the squares of the direction cosines of a line equals unity, we have

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

$$\text{or} \quad \left(\frac{d}{a/h}\right)^2 + \left(\frac{d}{b/k}\right)^2 + \left(\frac{d}{c/\ell}\right)^2 = 1$$

Upon rearranging, this becomes

$$d^2 = \frac{1}{\left(h^2/a^2 + k^2/b^2 + \ell^2/c^2\right)}$$

If the orthorhombic crystal is cubic, the basic interplanar distances are equal. In a simple cube, these distances are the lengths of the sides of a unit cell. We then have a = b = c and the equation above reduces to

$$d^2 = \frac{a^2}{\left(h^2 + k^2 + l^2\right)}$$

$$\text{or} \quad d = \frac{a}{\left(h^2 + k^2 + \ell^2\right)^{1/2}}$$

When this relation is substituted into Bragg equation we find that the first-order diffraction maxima in Bragg reflections are given by

$$\lambda = \frac{2a}{\left(h^2 + k^2 + \ell^2\right)^{1/2}} \sin\theta$$

If the wavelength λ of the radiation and the Bragg angle θ are known, then the equation above can be used to calculate the axial length of a side of a cubic crystal, and the Miller indices of the set of planes involved in giving a particular maximum.

Important Features of Miller Indices

There are many important features of Miller Indices some of which are mentioned below:

(i) Miller indices do not define a particular plane but a set of parallel planes.

(ii) All equally spaced parallel planes have the same index numbers h, k, *l*.

(iii) A plane parallel to one of the co-ordinate axes has an intercept of infinity along that axis, the corresponding index being zero.

(iv) A plane passing through the origin is defined in terms of a parallel plane having non-zero intercepts.

(v) It is only the ratio of the indices that is important. Thus the (422) planes are the same as the (211) planes.

Miller indices are especially useful for cubic crystals. Some of its desirable features for cubic system of crystals are

(i) The angle θ between two directions (u_1, v_1, w_1) and (u_2, v_2, w_2) can be easily calculated from the expression

$$\cos\theta = \frac{u_1u_2 + v_1v_2 + w_1w_2}{\left(u_1^2 + v_1^2 + w_1^2\right)^{1/2}\left(u_2^2 + v_2^2 + w_2^2\right)^{1/2}}$$

(ii) The distance d between two adjacent parallel planes with index number (hk*l*) is given by

$$d = \frac{a}{\left(h^2 + k^2 + \ell^2\right)^{1/2}}$$

Separation between lattice planes in simple, fcc and bcc cubic lattices

(a) **Simple cubic lattice (or cubic P lattice) :** In case of a simple cubic lattice a = b = c. The inter-planar separation for such a lattice is then given by

$$d_{hkl} = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$

Thus separation for (100) plane

$$d_{100} = \frac{a}{\sqrt{1^2 + 0^2 + 0^2}} = a$$

The separation for (110) plane

$$d_{110} = \frac{a}{\sqrt{1^2 + 1^2 + 0}} = \frac{a}{\sqrt{2}}$$

and for the (111) plane

$$d_{111} = \frac{a}{\sqrt{1^2 + 1^2 + 1^2}} = \frac{a}{\sqrt{3}}$$

The separation between successive (100), (110) and (111) planes are, therefore, a, $a/\sqrt{2}$ and $a/\sqrt{3}$ respectively. Hence, the ratio of their separation is

$$d_{100} : d_{110} : d_{111} = 1 : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{3}} = 1 : 0.71 : 0.58$$

(a) **fcc lattice :** Compared to simple cubic lattice (or cubic P) additional planes, shown by dotted line in Fig. 19.28 (a) & (b), arise half-way between (100) and (110) planes in case of fcc cubic lattice. These planes are drawn to include the atoms in face centres but do not

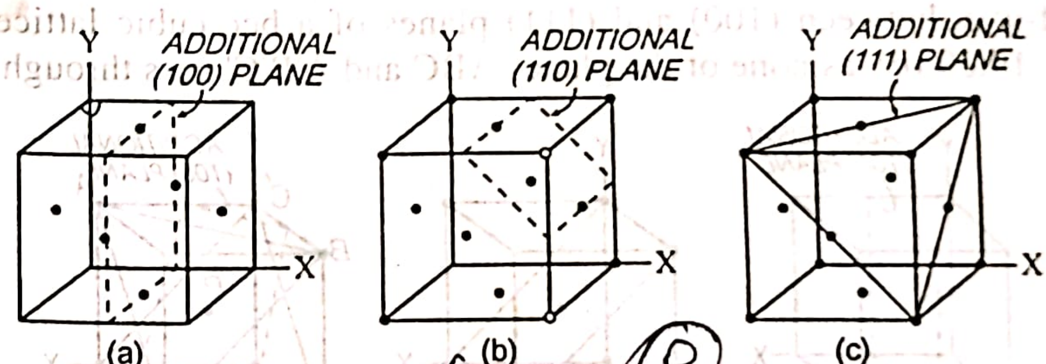


Fig. 19.28

pass through the corners of the unit cell. The (111) plane shown by solid line in Fig. 19.28 (c), passes through the corners and also contains three of the atoms in the face centres but no additional plane arises due to face centred lattice points as compared to simple cubic lattice. Thus the interplanar separation for fcc lattice is given by

(i) **for (100) plane**

$$d_{100} = \frac{1}{2}(d_{100})_P = \frac{1}{2}.a$$

(ii) **for (110) plane**

$$d_{110} = \frac{1}{2}(d_{110})_P = \frac{1}{2} . \frac{a}{\sqrt{2}} = \frac{a}{2\sqrt{2}}$$

(iii) **for (111) plane: since no additional plane arises due to face-centred lattice points,**

$$d_{111} = (d_{111})_P = \frac{a}{\sqrt{3}}$$

Thus for the face-centered cubic lattice, the largest separation is for the (111) plane while the smallest separation is for the (110) plane. The ratio of these separations are

$$d_{100} : d_{110} : d_{111} = \frac{a}{2} : \frac{a}{2\sqrt{2}} : \frac{a}{\sqrt{3}}$$

$$= \frac{1}{2} : \frac{1}{2\sqrt{2}} : \frac{1}{\sqrt{3}} = 0.87 : 0.61 : 1$$

(b) bcc lattice : When compared to simple cubic (or P) lattice, additional planes as shown by dotted line in Fig. 19.29 (a) & (c) arise half-way between (100) and (111) planes of a bcc cubic lattice. In Fig. 19.29 (c), as none of the planes ABC and A'B'C' pass through the

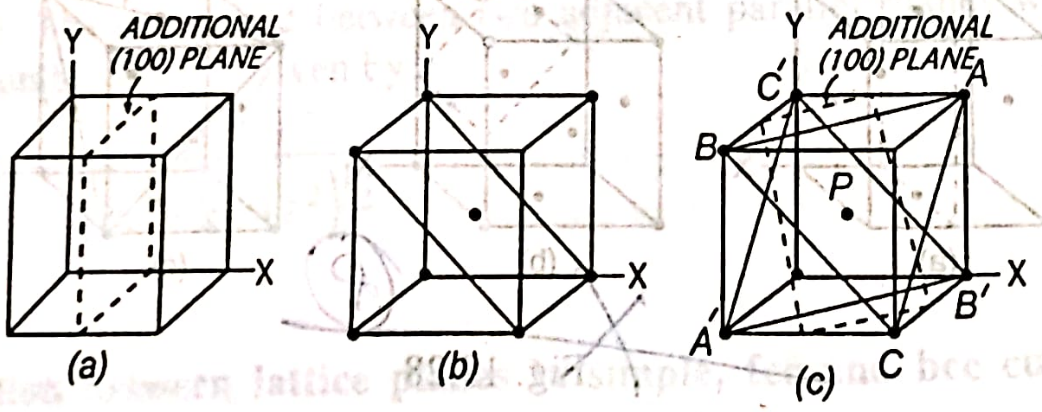


Fig. 19.29

central lattice point P, an additional plane has to be drawn between them. However no additional plane arises in case of (110) plane as compared to cubic P lattice. Thus the interplanar separation for bcc lattice is given by

(i) for (100) plane

$$d_{100} = \frac{1}{2}(d_{100})_P = \frac{a}{2}$$

(ii) for (110) plane

$$d_{110} = (d_{110})_P = \frac{a}{\sqrt{2}}$$

(iii) for (111) plane

$$d_{111} = \frac{1}{2}(d_{111})_P = \frac{1}{2} \cdot \frac{a}{\sqrt{3}} = \frac{a}{2\sqrt{3}}$$

Thus the largest separation in bcc lattice is for (110) plane while the smallest separation is for the (111) plane. The ratio of these separations is

$$d_{100} : d_{110} : d_{111} = \frac{a}{2} : \frac{a}{\sqrt{2}} : \frac{a}{2\sqrt{3}}$$

$$= \frac{1}{2} : \frac{1}{\sqrt{2}} : \frac{1}{2\sqrt{3}} = 0.71 : 1 : 0.41$$

Density of lattice points in a plane

In order to determine the number of lattice points in a lattice plane, consider N successive lattice planes, each having an area of cross-section A. Let d be the separation between successive lattice planes. Obviously, the volume of this part of the lattice is given by

$$V = N.A.d$$

If the volume of each unit cell is v, then the number of unit cells in a volume V of the lattice is given by

$$\frac{V}{v} = \frac{NAd}{v}$$

If n is the number of lattice points per unit cell, then the number of lattice point in a volume V is

n' = Number of unit cells x number of lattice points per unit cell

$$= \frac{NAd}{v} \times n$$

If ρ denotes the density of lattice points (number of lattice points per unit area) in these planes, then the total number of lattice points in the same part of the lattice (*i.e.*, in N planes), is A ρ N. This must be equal to n' *i.e.*, n' = A ρ N.

Equating the two expression for n'.

$$n' = A\rho N = \frac{NAd}{v}.n$$

$$\text{or,}\quad \rho = \frac{nd}{v}$$

In case of cubic, tetragonal, and orthorhombic systems of crystals, the volume of each unit cell v = a.b.c. Also for two primitive lattice in each of these systems, the number of lattice points per unit cell is one (*i.e.*, n=1). Therefore, in such cases

$$\rho = \frac{d}{abc}$$

***Example 19.1** Classify the following unit cells into proper system:*

(i) $a = 10.8\,\mathring{A},\ b = 9.47\,\mathring{A},\ c = 5.2\,\mathring{A}$
$\alpha = 41^\circ,\ \beta = 83^\circ,\ \gamma = 93^\circ$

(ii) $a = b = 10.73\,\mathring{A},\ c = 14.3\,\mathring{A}$
$\alpha = \beta = 90^\circ$ *and* $\gamma = 120^\circ$

Soln.

(i) Here a ≠ b ≠ c *i.e.*, the three primitives are unequal. In addition no two axes are at right angles to each other. Thus there is complete lack of symmetry which is characteristic of triclinic structure. Hence, this unit cell belongs to the triclinic system of crystals.

(ii) Here a = b ≠ c and α = β = 90° and γ = 120°. As seen in Table I, this unit cell belongs to hexagonal system of crystals.

***Example 19.2** The density of KCl crystal is 1.98 x 10^3 kg/m^3. Calculate the lattice constant and the distance between two adjacent atoms of the crystal. KCl has fcc lattice structure.*

Given: Atomic weight of potassium = 39.00 and that of chlorine = 35.55 and Avogadro's number = 6.02 x 10^{26} per kg-mol.

Soln.

Let a be the lattice constant.

Then mass in each unit cell = volume x density = $a^3\rho$ kg

Molecular weight (M) of KCl = (39+35.55) = 74.55

Since KCl belongs to fcc structure, there are 4 molecules per unit cell.

$$\therefore \text{ Mass in each unit cell } = \frac{n\,M}{N} \text{ kg}$$

where n is the number of molecules per unit cell and N the Avogadro's number

$$\therefore \quad a^3.\rho = \frac{n\,M}{N}$$

$$\text{or,} \quad a^3 = \frac{nM}{N\rho} = \frac{4 \times 74.55}{6.02 \times 10^{26} \times 1.98 \times 10^3}$$

$$= 25 \times 10^{-29} = 250 \times 10^{-30}$$

$$a = \left(250 \times 10^{-30}\right)^{1/3}$$

$$= 6.3 \times 10^{-10} = 6.3\,\mathring{A}$$

a is the lattice constant, *i.e.*, the distance between adjacent atoms of the same kind. Then the distance between two adjacent atoms *i.e.*, between potassium and chlorine is

$$d = \frac{a}{2} = \frac{6.3}{2}\mathring{A} = 3.15\,\mathring{A}$$

***Example 19.3** Calculate the lattice constant of potassium bromide from the following data:*

Atomic weight of potassium = 39.1

Atomic weight of bromine = 79.9

Density of potassium bromide = 2.7 x 10^3 kg/m^3

Avogadro's number = 6.02 x 10^{26} per kg-mode.

Potassium bromide has fcc crystal structure.

Soln.

Let "a" be the lattice constant and ρ the density of potassium bromide. Then

mass in each unit cell = $a^3 . \rho$

Also mass in each unit cell = $\frac{nM}{N}$

where n = No. of molecules per unit cell of potassium bromide = 4 (for fcc structure).

M = molecular weight of potassium bromide = (39.1 + 79.9) = 119.

N = Avogadro's number = 6.02 x 10^{26} per kg-mole.

$$\therefore \quad a^3\rho = \frac{nM}{N}$$

$$\text{or,} \quad a^3 = \frac{nM}{\rho N} = \frac{4 \times 119}{2.7 \times 10^3 \times 6.02 \times 10^{26}}$$

$$\text{or,} \quad a = \left[\frac{476}{16.254 \times 10^{29}}\right]^{1/3}$$

$$= 6.57 \times 10^{-10} \text{ m}$$

$$= 6.57 \text{ Å} .$$

Example 19.4 Calculate the lattice constant of cesium chloride which has a simple cubic lattice. The density of CsCl is 4 x 10^3 kg/m^3. Atomic weight of cesium = 132.9 and atomic weight of Cl = 35.5. Avogado's number = 6.02 x 10^{26} per kg-mole.

Soln.

As in the earlier examples, the lattice constant is given by

$$a = \left[\frac{n\,M}{N\rho}\right]^{1/3}$$

n for a simple cubic system is 1.

$$\therefore \quad a = \left[\frac{1 \times 168.4}{6.02 \times 10^{26} \times 4 \times 10^{3}}\right]^{1/3} \quad [M = 132.9 + 35.5 = 168.4]$$

$$= 4.12 \times 10^{-10}\ \text{m}$$

$$= 4.12\ \text{Å}.$$

Example 19.5 *Sodium is a bcc crystal. Its density is 9.6×10^2 kg/m^3 and atomic weight is 23. Calculate the lattice constant for a sodium crystal.*

Soln.

The lattice constant is given by

$$a = \left[\frac{n\,M}{N\rho}\right]^{1/3}$$

For a bcc crystal n = 2.

$N = 6.02 \times 10^{26}$ per kg-mode.

$$a = \left[\frac{2 \times 23}{6.02 \times 10^{26} \times 9.6 \times 10^{2}}\right]^{1/3}$$

$$= \left[0.8 \times 10^{-28}\right]^{1/3}$$

$$= 4.3 \times 10^{-10}\ \text{m}$$

$$= 4.3\ \text{Å}.$$

Example 19.6 *If density of NaCl is 2.163 gm/cm^3 and its molecular weight 58.45, find the spacing between planes parallel to the cubic lattice faces of the NaCl crystal. Given that Avogadro's number is 6.02×10^{23} molecules /g-mole.*

Soln.

From the definition of Avogadro's number (number of molecules in 1 gm-molecular weight of the substance), NaCl has 6.02×10^{23} molecules in 58.45gm. Hence

$$\text{No. of molecules/gm} = \frac{6.02 \times 10^{23}}{58.45} = 1.03 \times 10^{22}$$

$$\text{No. of molecules/cm}^3 = 1.03 \times 10^{22} \times 2.163$$

$$= 2.23 \times 10^{22}$$

Since NaCl has two atoms per molecule

$$\text{No. of atoms/cm}^3 = 2 \times 2.23 \times 10^{22} = 4.46 \times 10^{22}.$$

Crystal lattice of NaCl consists of a multitude of cubes with atoms at the corners, there will be $\left[4.46 \times 10^{22}\right]^{1/3} = 3.5 \times 10^{7}$ atoms in a row of 1 cm length. The spacing between atoms and hence between rows or planes is

$$d = \frac{1}{3.54 \times 10^{7}} = 2.83 \times 10^{-8}\text{cm} = 2.83\ \text{Å}$$

***Example 19.7** Copper has fcc structure and its atomic radius is 1.278 A.U. Calculate its density. Atomic weight of copper = 63.54.*

Soln.

As seen from Art. 19.10, the lattice constant for fcc structure is given by

$$a = 4r/\sqrt{2} = 4 \times 1.278/\sqrt{2} = 3.61\ \text{A.U.}$$

$$\text{Volume of unit cell} = a^3 = (3.61 \times 10^{-8})\text{cm}^3$$

$$= 47 \times 10^{-24}\ \text{cm}^3$$

If its density is ρ, then mass of unit cell

$$= 47 \times 10^{-24} \times \rho\ \text{gm}$$

Now let us find the mass of unit cell of copper in a different way. Each unit cell of fcc structure has 4 atoms. The mass of this 4 atoms obviously represent the mass of the cell. Since 6.02 x 10^{23} atoms of copper weigh 63.54 gm, the mass of 4 atoms

$$= \frac{4 \times 63.54}{6.02 \times 10^{23}} = 42.2 \times 10^{-23} \text{ gm}$$

Equating the two masses we have,

$$47 \times 10^{-24} \times \rho = 4.22 \times 10^{-23}$$

$$\text{or,} \quad \rho = \frac{42.2 \times 10^{-23}}{47 \times 10^{-24}} = 0.898 \times 10^{-23} \times 10^{24}$$

$$= 8.98 \text{ gm/cc}$$

Alternately, we know

$$a^3 \rho = \frac{n A}{N}$$

$$\text{or,} \quad \left(3.61 \times 10^{-8}\right)^3 \rho = \frac{4 \times 63.54}{6.02 \times 10^{23}}, \quad \text{or,} \quad \rho = 8.98 \text{ gm/cc}$$

Example 19.8 For a simple cubic lattice calculate

(i) the next neighbour distance, and

(ii) ratio of nearest neighbour distance to the next nearest neighbour distance

Soln.

(i) In a simple cubic lattice, the nearest neighbour distance $d_1 = a$ where a is the side of the unit cell. The next nearest neighbour lies at diagonally opposite end of the face of the unit cell. Therefore, the next neighbour distance

$$d_2 = \sqrt{a^2 + a^2} = a\sqrt{2}$$

(ii) The ratio of the nearest neighbour distance to the next nearest neighbour distance

$$d_1 : d_2 = a : a\sqrt{2}$$

$$= 1 : \sqrt{2}$$

***Example 19.9** In a crystal, a lattice plane cuts intercepts of a, 2b and 3c along the three axes, where a, b and c are primitive vectors of the unit cell. Determine the Miller indices of the given plane.*

Soln.

If the given plane cuts intercepts pa, qb and rc along the three axes, then

$$pa : qb : rc = a : 2b : 3c$$

or, $p : q : r = 1 : 2 : 3$

Then the Miller indices of the given plane

$$h:k:l=\frac{1}{p}:\frac{1}{q}:\frac{1}{r}=1:\frac{1}{2}:\frac{1}{3}$$

Multiplying the r.h.s by the L.C.M. 6

$$h : k : l = 6 : 3 : 2$$

Hence the Miller indices of the plane are (6 3 2)

***Example 19.10** In an ortho-rhombic crystal a lattice plane cuts intercepts of lengths 3a, -2b and 3c/2 along three axes. Deduce the Miller indices of the plane. a, b and c are primitive vectors of the unit cell.*

Soln.

For the given crystal, we have

$$pa:qb:rc=3a:(-2b):\frac{3c}{2}$$

or, $p:q:r=3:(-2):\frac{3}{2}$

The Miller indices of the given plane is

$$h:k:l=\frac{1}{p}:\frac{1}{q}:\frac{1}{r}$$

$$=\frac{1}{3}:-\frac{1}{2}:\frac{2}{3}$$

Multiplying the r.h.s. by 6, the L.C.M. of the denominator we get

$$h : k : l = 2 : (-3) : 4$$

Hence the Miller indices of the plane is $(2\bar{3}4)$.

Example 19.11 In a simple cubic crystal, find the ratio of intercepts on the three axes by (1 2 3) plane.

Soln.

If l_1, l_2 and l_3 are the lengths of the intercepts made by the plane on the three axes, then

$$l_1 : l_2 : l_3 = pa : qb : rc$$

where a, b, and c are the primitive vectors of the unit cell and the numerical parameters p, q, r are related to the Miller indices (h, k, l) of the plane by the relation

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r}$$

$$\text{or,} \quad p : q : r = \frac{1}{h} : \frac{1}{k} : \frac{1}{l}$$

Since the crystal in question is simple cubic

$$a = b = c$$

and for the given plane h = 1, k = 2, and l = 3.

$$\therefore \quad p : q : r = \frac{1}{1} : \frac{1}{2} : \frac{1}{3}$$

Multiplying the r.h.s. by 6, the LCM of the denominator

$$p : q : r = 6 : 3 : 2$$

We therefore have

$$l_1 : l_2 : l_3 = 6a : 3b : 2c$$

$$= 6a : 3a : 2a \qquad (\because a = b = c)$$

$$= 6 : 3 : 2$$

Example 19.12 The ratio of the axial units of a certain orthorhombic crystal is a : b : c = 0.424 : 1 : 0.367. Find the Miller indices of the crystal faces whose intercepts are

(i) 0.212 : 1 : 0.183

(ii) 0.848 : 1 : 0.732

(iii) 0.424 : ∝ : 0.123

Soln.

The axial units of the crystal are

a : b : c = 0.424 : 1 : 0.367

(i) pa = 0.212, or, p x 0.424 = 0.212

$$\therefore p = \frac{0.212}{0.424} = \frac{1}{2}$$

Similarly, qb = 1, or, q x 1 = 1

$\therefore$ q = 1

Also rc = 0.183, or, r x 0.367 = 0.183

$$\text{or,} \quad r = \frac{0.183}{0.367} = \frac{1}{2}$$

Hence the numerical parameters of the plane are $p : q : r = \frac{1}{2} : 1 : \frac{1}{2}$

Hence the Miller indices of the plane are

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} = 2 : 1 : 2 = (212)$$

(ii) The numerical parameters of this plane are

$$p = \frac{0.848}{0.424} = 2$$

$$q = \frac{1}{1} = 1$$

$$r = \frac{0.732}{0.367} = 2$$

The Miller indices are therefore

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} = \frac{1}{2} : \frac{1}{1} : \frac{1}{2}$$

Multiplying the r.h.s. by 2, the L.C.M. of the denominators, the Miller indices are,

$$h : k : l = 1 : 2 : 1 = (1\ 2\ 1)$$

(iii) The numerical parameters are

$$p = \frac{0.424}{0.424} = 1$$

$$q = \frac{\infty}{1} = \infty$$

$$r = \frac{0.123}{0.367} = \frac{1}{3}$$

The Miller indices are therefore

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} = \frac{1}{1} : \frac{1}{\infty} : \frac{1}{1/3} = 1 : 0 : 3 = (1 0 3)$$

Example 19.13 *In a crystal whose primitives are 1.2 Å, 1.8 Å, 2 Å, a plane (2 3 1) cuts an intercept of 1.2 Å, along the X-axis. Find the lengths of intercepts along Y-and Z-axis.*

Soln.

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} = 2 : 3 : 1$$

or $$p : q : r = \frac{1}{2} : \frac{1}{3} : \frac{1}{1}$$

Multiplying the r.h.s. by 6, the LCM

$$p:q:r=3:2:6$$

If l_1, l_2 and l_3 are the intercepts on the X-, Y-, and Z-axes respectively, we have

$$l_1 : l_2 : l_3 = pa : qb : rc$$

Here $a=1.2\,\mathring{A}, b=1.8\,\mathring{A}$ and $c=2\,\mathring{A}$

$$\therefore\ l_1:l_2:l_3=3\times(1.2):2\times(1.8):6\times(2)$$

$$=3.6:3.6:12$$

Now $\dfrac{l_2}{l_1}=\dfrac{3.6}{3.6}$ $\quad\therefore\ l_2=\dfrac{3.6}{3.6}\times1.2\,\mathring{A}\ \ (l_1=1.2\,\mathring{A})$

$$=1.2\,\mathring{A}$$

Again $\dfrac{l_3}{l_1}=\dfrac{12}{3.6}$ $\quad\therefore\ l_3=\dfrac{12}{3.6}\times1.2\,\mathring{A}$

$$=4\,\mathring{A}$$

Thus the intercepts along the Y-and Z-axes are $1.2\,\mathring{A}$ and $4\,\mathring{A}$ respectively.

Example 19.14 The lattice constant of a crystal are 1.21 $\mathring{A}$, 1.84 $\mathring{A}$ and 1.98 $\mathring{A}$. A plane whose Miller indices are (1, -2, -3) cuts intercept of – 1.32 $\mathring{A}$ along the Z-axis. What are the lengths of the intercepts along Y – and Z – axes.

Soln.

$$h:k:l=\frac{1}{p}:\frac{1}{q}:\frac{1}{r}=1:-2:-3$$

or, $\quad p:q:r=\dfrac{1}{1}:\dfrac{1}{-2}:\dfrac{1}{-3}$

Multiplying the right hand side by the lowest common denominator 6, we have

$$p:q:r=6:-3:-2$$

If l_1, l_2 and l_3 are the lengths of the intercepts made by the plane in question along X-, Y-, and Z- axes respectively; then

$$l_1 : l_2 : l_3 = pa : qb : rc$$

$$= 6\,(1.21) : -3\,(1.84) : -2\,(1.98)$$

Here $a = 1.21\,\mathring{A}, b = 1.84\,\mathring{A}$ and $c = 1.98\,\mathring{A}$

or, $l_1 : l_2 : l_3 = 7.26 : -5.52 : -3.96$

But $l_3 = -1.32\,\mathring{A}$

Hence from $\frac{l_2}{l_3} = \frac{-5.52}{-3.96}$ we have $l_2 = \frac{-5.52}{-3.96} \times (-1.32)$

$$= -1.84\,\mathring{A}$$

Again from $\frac{\ell_1}{\ell_3} = \frac{7.26}{-3.96}$

we have $l_1 = \frac{7.26}{-3.96} \times (-1.32)$

$$= 2.42\,\mathring{A}$$

Thus the intercepts along X- and Y- axis are $2.42\,\mathring{A}$ $-1.84\,\mathring{A}$ respectively.

Example 19.15 The lattice constants of a simple lattice is a. Find the lattice spacings between (111), (112) and (113) lattice planes.

Soln.

The lattice spacing for a given set of parallel planes (hk$l$) is given by

$$d_{hkl} = \frac{1}{\left[\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}\right]^{1/2}}$$

For (111) lattice planes

h = 1, k = 1 and l = 1 and since the crystal is simple cubic, a = b = c

$$d_{111} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{1^2}{a^2}+\frac{1^2}{a^2}\right]^{1/2}} = \frac{1}{(3/a^2)^{1/2}} = \frac{a}{\sqrt{3}}$$

Similarly for the (112) lattice planes, h = 1, k = 1, l = 2

or, $$d_{112} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{1^2}{a^2}+\frac{2^2}{a^2}\right]^{1/2}} = \frac{1}{(6/a^2)^{1/2}} = \frac{a}{\sqrt{6}}$$

and for the (113) planes

$$d_{113} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{1^2}{a^2}+\frac{3^2}{a^2}\right]^{1/2}} = \frac{1}{(11/a^2)^{1/2}} = \frac{a}{\sqrt{11}}$$

Example 19.16 For a simple cubic lattice of lattice parameter 2.04 Å, *calculate the spacing of lattice planes (2 1 2).*

Soln.

$$d_{hkl} = \frac{1}{\left[\frac{h^2}{a^2}+\frac{k^2}{a^2}+\frac{l^2}{a^2}\right]^{1/2}}$$

Here h = 2, k = 1, l = 2

and a = b = c = 2.04 Å

$$\therefore\ d_{212} = \frac{1}{\left[\frac{2^2}{(2.04)^2}+\frac{1^2}{(2.04)^2}+\frac{2^2}{(2.04)^2}\right]^{1/2}}$$

$$= \frac{1}{\left[\frac{4}{4.16}+\frac{1}{4.16}+\frac{4}{4.16}\right]^{1/2}}$$

$$= \frac{1}{(9/4.16)^{1/2}} = \frac{2.04}{3} = 0.68\ \mathring{A}$$

Example 19.17 In a tetragonal lattice a = b = 2.12 Å, c = 1.81 Å. Find (i) the lattice separation between (110) planes, (ii) the density of lattice points in (110) planes.

Soln.

(i) The lattice separation is given by

$$d_{hkl} = \frac{1}{\left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right]^{1/2}}$$

Here h = 1, k = 1 and $l = 0$

and a = b = 2.12 Å and c = 1.81 Å

$$\therefore\ d_{110} = \frac{1}{\left[\frac{1^2}{(2.12)^2}+\frac{1^2}{(2.12)^2}+\frac{0}{(1.81)^2}\right]^{1/2}}$$

$$= \frac{1}{\left[\frac{1^2}{4.49}+\frac{1^2}{4.49}+0\right]^{1/2}} = \frac{1}{(2/4.49)^{2}} = \frac{1}{(0.445)^{1/2}}$$

$$= 1.5\ \mathring{A}$$

(ii) The density of lattice points is given by

$$\rho = \frac{nd}{v}$$

where n = number of lattice points per unit cell

d = lattice separation

v = volume of unit cell

For a tetragonal system, n = 1 and v = a.b.c

$$\therefore\ \rho = \frac{1.d}{v} = \frac{1.5\overset{o}{A}}{\left(2.12\overset{o}{A}\right)\times\left(2.12\overset{o}{A}\right)\times\left(1.81\overset{o}{A}\right)}$$

$$= \frac{1.51\times10^{-10}}{8.1348\times10^{-30}} = 0.184\times10^{20} = 1.84\times10^{19}\ \text{atoms/m}^3$$

***Example 19.18** Show that in a simple cubic lattice the separation between the successive lattice planes (100), (110) and (111) are in the ratio 1: 0.71 : 0.58.*

Soln.

For simple cubic lattice a = b = c

$$d_{100} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{0}{a^2}+\frac{0}{a^2}\right]^{1/2}} = \frac{1}{\left(1/a^2\right)^{1/2}} = a$$

$$d_{110} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{1}{a^2}+\frac{0}{a^2}\right]^{1/2}} = \frac{1}{\left(2/a^2\right)^{1/2}} = \frac{a}{\sqrt{2}}$$

$$d_{111} = \frac{1}{\left[\frac{1^2}{a^2}+\frac{1^2}{a^2}+\frac{1^2}{a^2}\right]^{1/2}} = \frac{1}{\left(3/a^2\right)^{1/2}} = \frac{a}{\sqrt{3}}$$

$$\therefore\ d_{100} : d_{110} : d_{111} = a : \frac{a}{\sqrt{2}} : \frac{a}{\sqrt{3}} = 1 : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{3}}$$

$$= 1 : 070 : 058.$$

***Example 19.19** Deduce the density of lattice points in (100), (110) and (111) planes in NaCl cubic lattice which belongs to fcc lattice. Given: lattice parameter = 5.63, number of lattice points per unit cell = 4.*

Soln.

The density of lattice points is given by

$$\rho = \frac{nd}{v}$$

where n = no. of lattice points per unit cell

d = separation between lattice planes

v = volume of each unit cell (a^3)

for fcc crystal structure n = 4

$$d_{hkl} = \frac{1}{\left[\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}\right]^{1/2}}$$

For (100) plane, $d_{100} = \frac{a}{2}$

$$\therefore \rho_{100} = \frac{n \cdot d_{100}}{v} = \frac{4 \times a}{2 \times a^3}$$

$$= \frac{4 \times 5.63 \mathring{A}}{2 \times \left(5.63 \mathring{A}\right)^3} = \frac{4 \times 5.63 \times 10^{-10} \text{ m}}{2 \times (5.63)^3 \times 10^{-30} \text{ m}^3}$$

$$= 0.0631 \times 10^{20} \text{ atom/m}^3$$

$$= 6.31 \times 10^{18} \text{ atoms/m}^3$$

For (110) plane, $d_{110} = \frac{a}{2\sqrt{2}}$

$$\therefore 110 = \frac{nd_{110}}{a^3} = \frac{4 \times 5.63 \times 10^{-10}}{2\sqrt{2} \times (5.63)^3 \times 10^{-30}}$$

$$= 0.446 \times 10^{20}$$

$$= 4.46 \times 10^{18} \text{ atoms / m}^3$$

For (111) plane, $d_{111} = \frac{a}{\sqrt{3}}$

$$\therefore \ 111 = \frac{nd_{111}}{a^3} = \frac{4 \times a}{\sqrt{3} \times a^3}$$

$$= \frac{4 \times 5.63 \times 10^{-10}}{1.732 \times (5.63)^3 \times 10^{-30}}$$

$$= 0.0728 \times 10^{20}$$

$$= 7.28 \times 10^{18} \text{ atoms /m}^3$$

V.V.I

Example 19.20 Find the Miller indices of a set of parallel planes which make intercept in the ratio 3a : 4b on the X-and Y-axes and are parallel to the Z-axis, a, b and c being the primitive vectors of the lattice.

Also calculate the inter-planer distance of these planes taking the lattice to be a cubic lattice with $a = b = c = 2.0 \times 10^{-10}$ m.

Soln.

As the set of parallel planes is parallel to the Z-axis, their intercept on the Z-axis is infinite. Also the intercepts on X- and Y-axes are in the ratio 3a : 4b. Thus

$$pa : qb : rc = 3a : 4b : \propto c$$

or $\quad p : q : r = 3 : 4 : \propto$

Hence the Miller indices of the plane

$$h : k : \ell = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} = \frac{1}{3} : \frac{1}{4} : \frac{1}{\propto} = \frac{1}{3} : \frac{1}{4} : 0$$

Multiplying the right hand side of the equation by the lowest common denominator, we get

$$h : k : \ell = 4 : 3 : 0$$

Hence the Miller indices of the plane are (430).

The inter-planer distance for a given set of parallel planes ($h\,k\,\ell$) is given by

$$d_{hkl} = \frac{1}{\left[\frac{h^2}{a^2} + \frac{k^2}{a^2} + \frac{l^2}{a^2}\right]^{1/2}}$$

Since the lattice in question is cubic, a = b = c. Also h = 4, k = 3 and $l = 0$

$$\therefore \; d_{430} = \frac{1}{\left[\frac{4^2}{a^2} + \frac{3^2}{a^2} + \frac{0}{a^2}\right]^{1/2}}$$

$$= \frac{1}{[25/a^2]^{1/2}} = \frac{a}{5}$$

$$= \frac{2.0 \times 10^{-10}}{5} \quad [\because a = 2.0 \times 10^{-10}]$$

$$= 0.4 \times 10^{-10} \quad = 0.4 \text{Å}$$

19.14 Crystal Defects:

In our discussions so far only ideal crystals have been considered. An ideal crystal is a crystal in which each atom occupies a definite location in a regular array – each atom has identical surroundings. Actual crystals are far from meeting this specification. For one thing, the atoms in a crystal are in continual thermal vibration. However, these vibrations represent imperfections only in a limited sense, since the equilibrium positions of the atom can still be described in terms of a perfect lattice. More drastic are the actual defects in the structure of a crystal. These defects may be a missing atom, an atom in interstitial positions, irregularities in the periodicity of the lattice along certain lines, and so forth. These defects can be reduced considerably but can never be eliminated entirely. What is more, in some situations these defects are highly desirable. For instance, the donor and acceptor impurities are essential to the operation of transistors. The mechanical strength of a solid is largely determined by the nature and concentration of defects in its structure. The most important types of crystal defects are (1) point defects, (2) line defects, and (3) surface defects.

Point Defects:

The simplest type of crystal imperfection is the localized point defect. During crystallization process, all sorts of impurities attach themselves to a crystal. In addition, impurities are usually deliberately added to pure crystals to modify their properties. For example, the conductivity of germanium crystal is enormously enhanced by the addition of pentavalent and trivalent impurities. Point defects are of three basic kinds: the substitutional, the interstitial and the vacancy.

Substitutional:

An impurity is *substitutional* if it occupies a lattice site from which a host atom has been expelled [Fig. 19.30(a)]. The impurity is interstitial if it occupies a position between the host atoms [Fig. 19.30(b)]. If the impurity atom has roughly the same size and valency as the atoms, then the substitutional impurity is created. It is very difficult to insert such an atom in the interstitial position as considerable

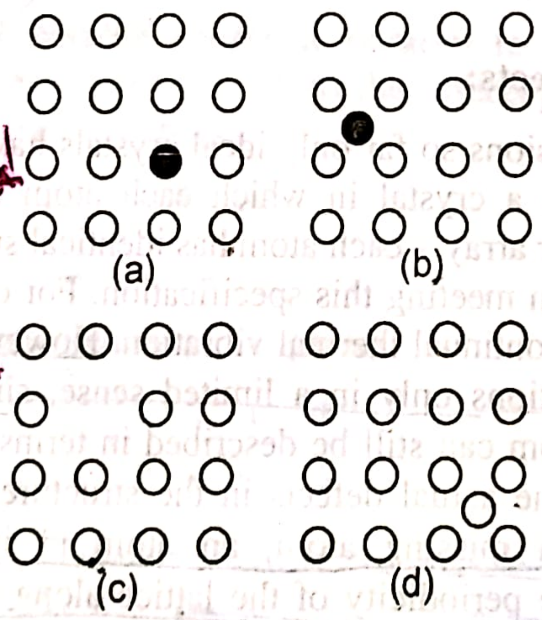


Fig. 19.30

energy will be needed to squeeze the neighbouring atoms more tightly despite their mutual repulsions in order to make room. Only impurity atoms smaller in size than the host atoms in a crystal are able to fit more readily into interstitial locations and may be present

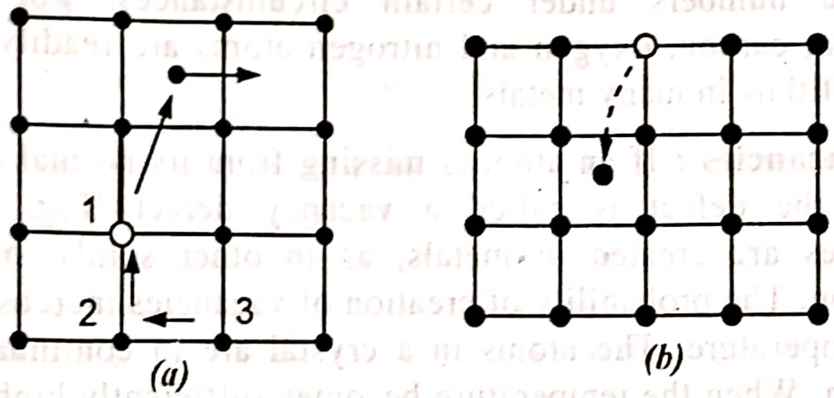
in large numbers under certain circumstances. For example, hydrogen, carbon, oxygen and nitrogen atoms are readily absorbed as interstitials in many metals.

Vacancies : If an atom is missing from its normal site in the matrix, the defect is called a vacancy defect [Fig. 19.30(c)]. Vacancies are created in metals, as in other solids, by thermal excitation. The probability of creation of vacancies increases rapidly with temperature. The atoms in a crystal are in continual thermal vibration. When the temperature becomes sufficiently high, some of the vibrating atoms acquire enough energy to leave the site completely. When the atom leaves, the region surrounding the vacancy is distorted. The distortion is the result of the *relaxation* of the lattice, as it were, in order to partially fill the void left by the atom. The distortion contributes further to the irregularity of the lattice in the immediate neighbourhood of the vacancy.

Interstitials : This defect is due to the presence of a regular atom in an interstitial position [Fig. 19.30(d)]. A considerable amount of energy is needed to pull out an atom from its regular position and place it in an interstitial position. The defect is created thermally only at high temperatures, near the melting point of the solid.

Frenkel and Schottky defects :

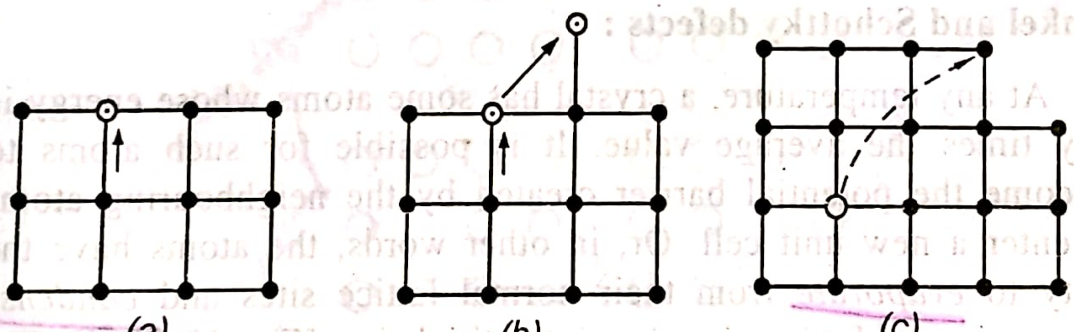
At any temperature, a crystal has some atoms whose energy is many times the average value. It is possible for such atoms to overcome the potential barrier created by the neighbouring atoms and enter a new unit cell. Or, in other words, the atoms have the ability to *evaporate* from their normal lattice sites and *condense* between normal lattice sites *i.e.*, interstitial sites [Fig. 19.31(a)]. The result is the simultaneous creation of a vacancy and an interstitial atom. The pair of vacancy and the interstitial atom, thus created, is called the *Frenkel defect*. The same defect may also be created when a surface atom diffuses into interstitial sites, thereby creating a pair of vacancy and interstitial [Fig. 19.31(b)] atom.



(a) (b)

Fig. 19.31

Atoms can evaporate completely or partially from the surface of a crystal. In case of complete evaporation, an atom leaves the surface of the crystal and moves into the vapour [Fig. 19.32(a)]. In case of partial evaporation, the atom leaves the surface of the crystal and moves into the space above the surface [Fig. 19.32(b)]. Whether the evaporation is complete or partial, a vacancy is created in each case. And as a deeper lying atom moves in to fill up the vacancy, another vacancy is created below the surface, *i.e.*, the vacancy is pulled inside the crystal and diffuses throughout its bulk. The vacancies created in this manner are called *Schottky* defect. Schottky effect may also be created by the transition of atoms from inside the crystal into its surface.



(a) (b) (c)

Fig. 19.32

Point defects in excess of the number in thermal equilibrium may be created in a number of ways. The simplest method is to heat a solid to a high temperature so that many defects are present and then to cool it quickly enough to retain some of these defects at the final temperature. Of much importance is the production of defects by particle radiation. For instance, high energy neutrons in a nuclear

reactor produce these defects by simply knocking atoms out of their normal locations. The properties of nearly all solids are severely affected by the intense bombardment by neutrons. Radiation damage leads to a reduction in ductility (embrittlement) in most metals, often accompanied by an increase in the yield strength.

The existence of point defects in a crystal makes it possible for the diffusion of atoms within the crystal to occur. When a vacancy is present, diffusion occurs by the jumping of an adjacent atom to fill it in a sort of relay race. As atom 2 moves into vacancy 1, the vacancy appears to move into point 2. As atom 3 moves into this vacancy, the latter appears to move into point 3 and so on.

In effect, the vacancy moves from place to place through the crystal. When an interstitial is present, diffusion occurs as it migrates from one interstitial location to a neighboring one. In both cases, each change in position requires a certain amount of extra energy, in order for the moving atom to squeeze past the atoms that project in its path. Thus the diffusion in a solid, as we might expect, is strongly temperature dependent, increasing from a usually negligible rate at room temperature to one not far from that characteristic of the corresponding liquid near the melting point.

Diffusion in an ionic crystal enables it to conduct electric current, since the moving defects are either charged interstitials or vacancies that behave as though they carry charges opposite to the missing ions. In fact, much detailed information on point defects has been obtained from the study of the electrical conductivity of the ionic crystals. These informations would have been difficult to obtain in any other way.

Dislocations – line defects

Line defects in a crystal are called *dislocations*. A dislocation is a line of atoms with different co-ordinations from those of the other atoms in the crystal. There are two basic types of dislocation – the *edge dislocation* and the *screw dislocation*. The edge dislocation is illustrated in Fig. 19.33, which can be visualized in terms of the removal of part of one layer of atoms and the subsequent accommodation of the array to the defect. The crystal structure in the immediate neighbourhood of the dislocation, indicated by the symbol ⊥,

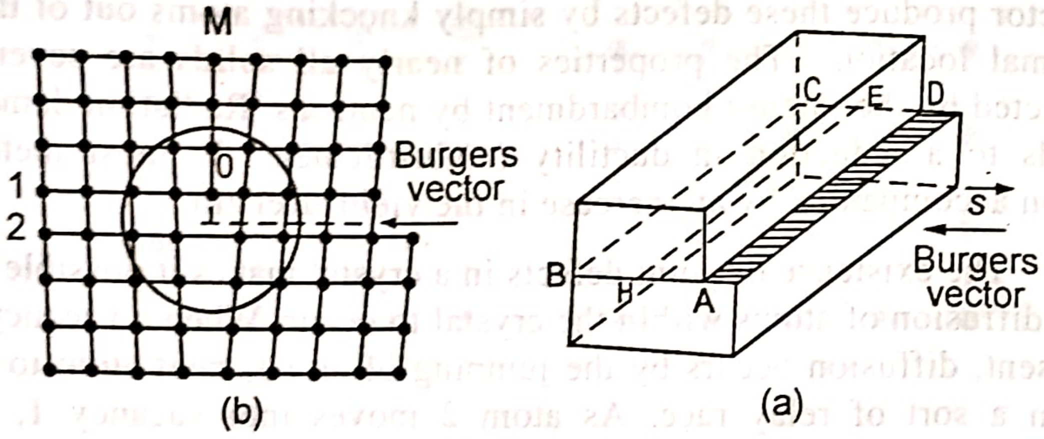


Fig. 19.33

is severely distorted. Far away from the dislocation, the crystal regains its regularity.

The other kind of dislocation is the screw dislocation which is illustrated in Fig. 19.34. The formation of screw dislocation may be visualized by imagining that a cut ABCD is made a certain distance into a crystal. The left side is then slipped up past the right side. Lines on top represent vertical atomic planes. Shaded area ABB' indicates the region

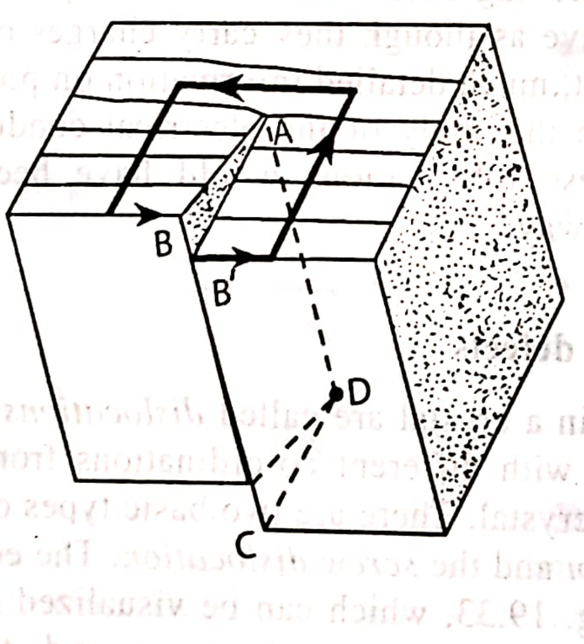


Fig. 19.34

of slippage. The points B and B' were coincident before the dislocation was created. The dislocation is represented by line AD which lies at the end of the step BAB' created by the slip. The atomic planes around the dislocation actually spiral around it as indicated by the arrows. This accounts for the name screw dislocation. Actual dislocations in crystals are usually combination of the edge and screw varieties.

19.15 Burgers vector

A dislocation can be described by means of a closed loop surrounding the dislocation line. The loop which is also referred to as *Burgers* circuit, is formed by proceeding through the undisturbed region surrounding a dislocation in steps which are integral multiples of a lattice translation. The loop is completed by going an equal number of translations in both positive and negative sense in a plane normal to the dislocation line. The amount by which the circuit fails to close upon itself is the Burger's vector **s** and is given by

$$\mathbf{s} = n_a\mathbf{a} + n_b\mathbf{b} + n_c\mathbf{c}$$

where n_a, n_b or n_c are either zero or integers and **a**, **b** and **c** are the three primitive lattice transformations. A zero value of **s** indicates that there is no internal slip in the crystal. Fig. 19.35 illustrates a Burger's circuit for a screw dislocation. It can be seen that starting at some lattice point at the front of the crystal, the loop fails to close on itself by one unit translation parallel to the dislocation line.

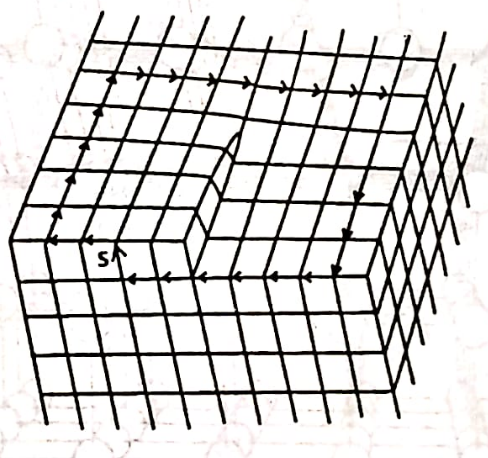
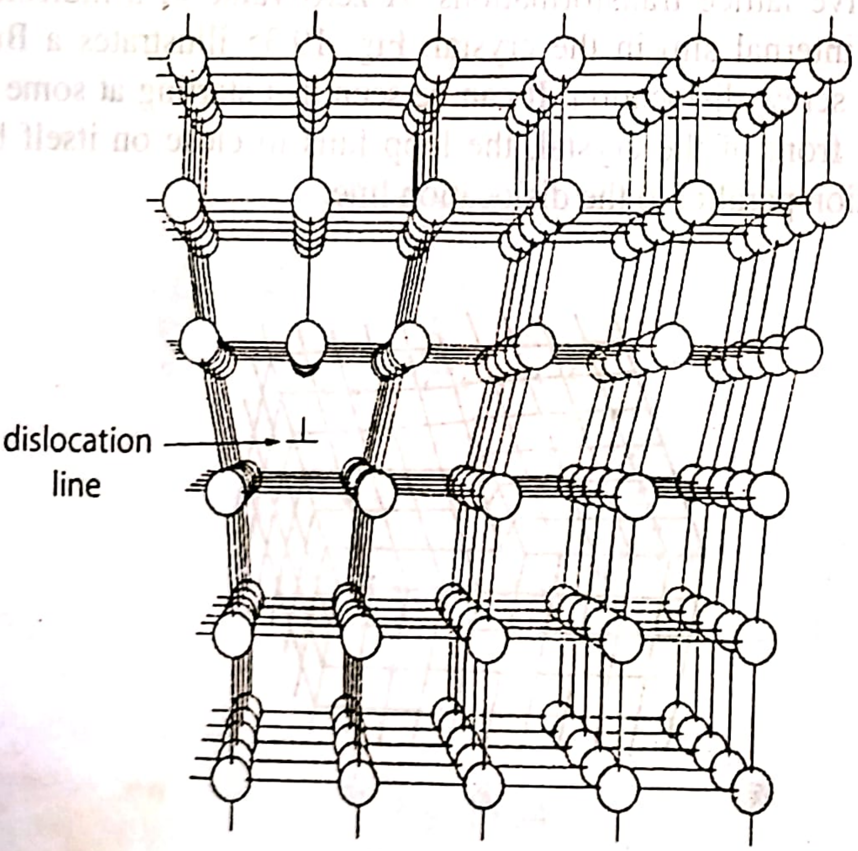


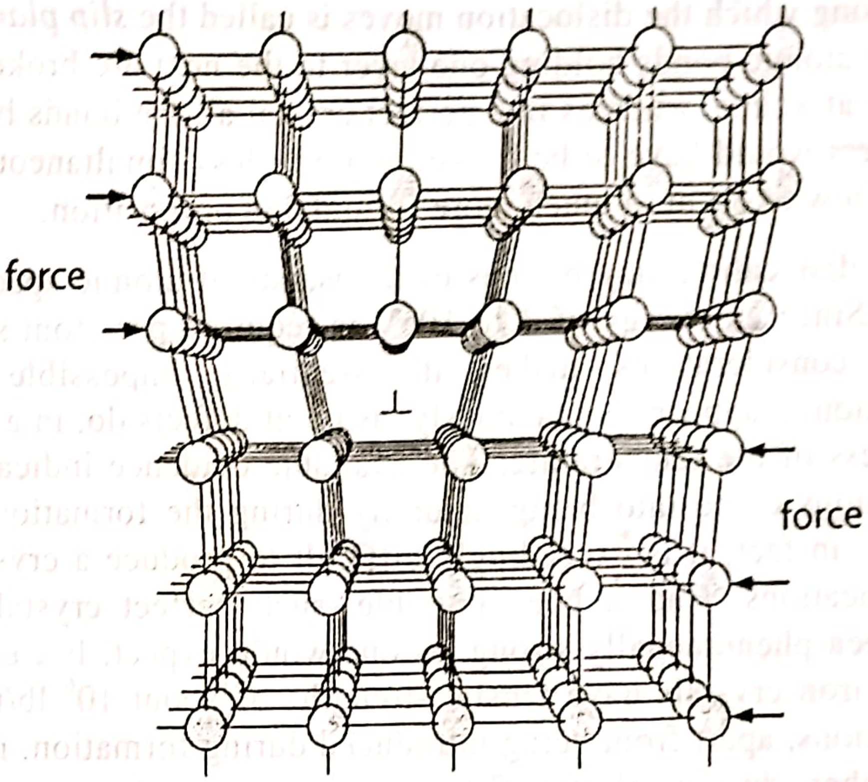
Fig. 19.35

This is the Burgers vector s. The Burgers vector of a screw dislocation is parallel to the dislocation line. It may be noted that if the loop is continued, it will describe a spiral path around the Burgers dislocation similar to the thread of a screw.

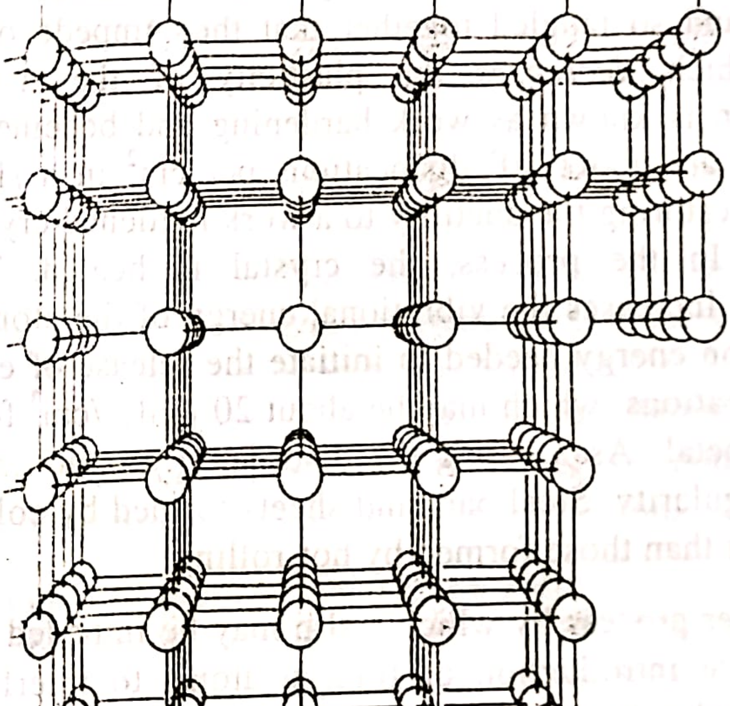
Dislocations provide an explanation for the plastic behaviour of solids. The elastic response of a solid can be readily interpreted in terms of the bonding forces within it, which act like Hooke's law restoring forces for small displacements from the equilibrium configuration. But the plastic behaviour of solids cannot be accounted for by this direct approach, since calculations of the force required to slide one layer of atoms in a crystal past another yield figures which is approximately 10^3 times heigher than those actually observed. But the presence of dislocation makes it possible to understand why solids are only about 0.1 percent as strong as they ought to be on the basis of perfect crystal structure. Fig. 19.36 illustrates how a crystal having an edge dislocation can be permanently deformed by a relatively modest shearing stress. When the shear is applied, the line of atoms which is below and to the right of the edge dislocation shifts its bonds to the line directly above it. The dislocation thus moves one atom spacing to the right. The process is repeated until the dislocation moves to the edge of the crystal when the deformation becomes permanent. The entire process is called *slip*, and the



(a)



(b)



(c)

Fig. 19.36

plane along which the dislocation moves is called the *slip plane*. In a slip, the atomic bonds holding one layer to the next are broken only one line at a time, whereas in a perfect crystal all the bonds between two layers would have to be broken more or less simultaneously for plastic flow to occur, a much more formidable proposition.

A dislocation may be tens of thousands of atomic spacings in length. Since an energy of 3 to 10eV is required per atom spacing, energy considerations makes it essentially impossible for a dislocation to appear spontaneously, as point defects do, in a crystal, regardless of the temperature. The available evidence indicates that dislocations come into being naturally during the formation of the crystal - in fact, it is exceedingly difficult to produce a crystal free of dislocations. Had it been possible, such perfect crystal would have been phenomenally strong, as one would expect. For example, perfect iron crystals have tensile strengths of about 10^6 lb/in^2. The dislocations, apart from being introduced during formation, multiply in number during plastic flow in a variety of more or less complicated ways. This continued deformation of a solid increases its dislocation content. Eventually the dislocation becomes so numerous and so tangled together that they impede one another's motion, which *decreases* the plasticity of the material. This phenomenon is known as work hardening and becomes prominent when there are about 10^8 dislocations per cm^2 in a slip line. The process of restoring the ductility to a work hardened crystal is called *annealing*. In the process, the crystal is heated. The rise in temperature increases the vibrational energy of the atoms providing the activation energy needed to initiate the release of energy stored in the dislocations, which may be about 20 joules/cm^3 for a severely deformed metal. As a result the disordered atomic array tends to return to regularity. Steel bars and sheets formed by cold-rolling are much harder than those formed by hot-rolling.

Another process by which a slip may be impeded in a metal is the deliberate introduction of foreign atoms to interfere with the motion of dislocations. Thus the addition of small amounts of such other elements of carbon, chromium, manganese, and tungsten to iron converts it into the vastly stronger steel by reducing slip.

19.16 Plane (or surface) Defects

In plane defects, the crystalline irregularity extends in two dimensions. The most common type of plane defects in any crystal is its surface itself. The atoms at the surface of a crystal are not surrounded by as many numbers of atoms as those inside the crystal. The energy inside the crystal is therefore less than at the surface. For the same crystal structure, the co-ordination number of an atom inside the crystal is different from that of an atom at the surface. Thus, the surface of a crystal is a defect on its own. This surface has a decisive effect on the properties of samples such as thin films and fibres. Some representative surface or plane defects are mentioned below.

(i) Stacking fault:

Close packing can give rise to plane defects. Suppose that in a closest packing the n^{th} layer is an A layer and the $(n + 1)^{th}$ layer is supposed to be a B layer but because of a *mistake* in the stacking sequence it is a C layer instead. A stacking fault is said to have been introduced between the n^{th} and $(n + 1)^{th}$ layer.

(ii) Grain boundaries:

The imperfections discussed above can exist in individual crystals. It is therefore vary difficult to grow single crystals in which the lattice planes are arranged in a perfectly regular manner. Most common materials consist of many small single crystals stuck together in random manner. The boundary between two crystals is called a *grain boundary*. A grain boundary must therefore have a structure that somehow conforms to the structures and orientations of both the crystals or grains. Obviously, the grain boundary forms a discontinuity in the periodicity of the lattice of either grain or crystal and is, therefore a type of lattice imperfection.

In the region of grain boundary, the atoms are not arranged in proper order and are always under pull by the two adjacent grains. As a consequence, the grain boundary atom cannot join either of the crystals and therefore, takes up a compromising position. As the grain boundary atoms are obviously not surrounded by as many atoms as those inside the crystal, the energy of the grain boundaries is more than that of the grains. Because the opposing forces experienced by grain boundary atoms is restricted to short distances

only, the thickness of grain boundary surface is only a few atomic diameters. If the misorientation between the two grains is more than 10-15°, the *grain boundary* is called a high-angle grain boundary. If the angle between the orientations of the two grains is less than 10°, the boundary is known as a low-angle grain boundary.

Some of the other important crystal imperfections are discussed below.

(i) Electronic imperfections

This defect arises due to error in charge or energy distribution in solids and plays an important part in the electrical and magnetic properties of materials.

(ii) Excitation states of crystals

Excitation states are quantized. These states can be regarded as imperfections as these cause deviations from perfect crystal symmetry. Some examples are

(a) phonons and magnons which are quantized lattice vibrations and spin waves respectively.

(b) conduction electrons and holes which are excited thermally from filled bands impurity levels.

(c) excitons which are quantized electron hole pairs.

(d) quantized plasma waves.

(iii) Transient imperfections

These imperfections are induced into the crystal due to external sources *e.g.*,

(a) photons, which are ordinary quantized electromagnetic waves.

(b) high energy charged particles like electrons, protons, mesons, ions, etc.

(c) high energy uncharged particles such as neutrons and neutral atoms.

Plastic deformations

Elastic properties are reversible. If a crystal is elastically deformed under the influence of an applied stress, it will return to its

original form when the stress is removed. However, if the applied stress is sufficiently large, a certain degree of deformation remains even after the removal of the stress. The crystal is said to have been plastically deformed. This is referred to as the plasticity of the crystal. Plastic deformations are, therefore, irreversible deformations.

Plastic deformation takes place in a crystal due to the sliding of one part of the crystal with respect to the other. This results in a slight increase in length of the crystal. This is illustrated in Fig. 19.37.

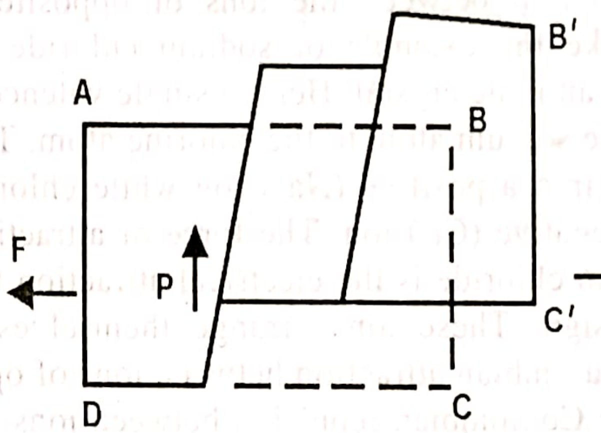


Fig. 19.37

A tension FF is applied to the crystal ABCD resulting in a slight increase in its length. The process of sliding is called a *slip*. The direction and plane along which the sliding takes place are called the slip direction as shown by the arrow P and the slip plane respectively.

19.17 Types of bonds in solids

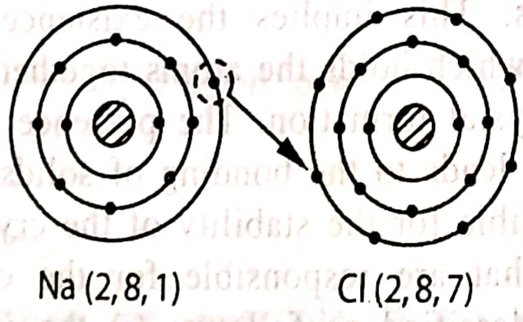
Solids are stable structure, *e.g.*, a crystal of sodium chloride is more stable than a collection of free sodium and chlorine atoms. Similarly, a Germanium crystal is more stable than a collection of Germanium atoms. This implies the existence of an attractive *interatomic* force which holds the atoms together. This is the force responsible for crystal formation. The presence of these attractive interatomic forces leads to the bonding of solids. It is these bonds which are responsible for the stability of the crystal. The different types of bonds that are responsible for the cohesive forces in crystals can be classified as follows: (i) the *ionic bond* (ii) the

covalent bond (iii) the *metallic bond,* (iv) the *molecular or Van der Wall's bond* and (v) the *hydrogen bond.*

The attractive interatomic force which holds the atoms and molecules together in solids are almost entirely electrical in nature. The distinctive differences between the various types of crystal bonding may be attributed to the different types of electron distribution around the atoms and molecules.

(i) Ionic bond : In ionic bond, the force between atoms is the electrical attraction between the ions of opposite signs. For example let us take the example of sodium chloride which is a typical example of an ionic crystal. Here, a single valence electron is transferred from the sodium atom to the chlorine atom. Thus sodium atom is converted into a positive (Na^+) ion while chlorine atom is converted into a negative (Cl^-) ion. The force of attraction between the atoms of sodium chloride is the electrical attraction between the ions of opposite sign. These ions arrange themselves in such a manner that the Coulombian attraction between ions of opposite sign is stronger than the Coulombian repulsion between ions of the same sign. This type of binding between ions of some crystals is referred to as the *ionic* bond. The ionic binding is valid upto a particular distance between the opposite ions. For smaller distances, the repulsive force become significant and the crystal structure is maintained stable.

Ionic crystals are formed by elements of group I and VII. The cohesive energy (binding energy) of an ionic crystal is equal to the energy required to transform the crystal into a system of isolated ions. The higher the cohesive energy, the more stable the crystal is. Thus ionic crystals are liable to break under stress. As there is no free electrons in an ionic crystal, these crystals are poor conductor of electricity. But due to their high cohesive energy, the ionic crystals



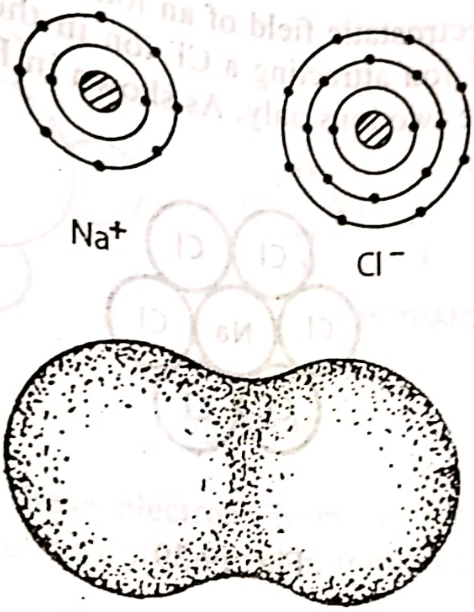


Fig. 19.38

have high melting point. For example, the melting points of Na_2SO_4, an ionic crystal, is as high as 1157°K. An example of ionic bonding is illustrated in Fig. 19.38.

Mainly formed in inorganic compounds like NaCl, KOH etc. and never in pure elements.

	Ionic Binding
Constituent Particles	Positive and negative ions arranged systematically.
Binding forces	Strong electrostatic attraction
Properties	Brittle, soft, poor conductors of heat and electricity, high melting point and high latent heat of fusion. Soluble in solvents like water (H_2O) and liquid ammonia (NH_3). Dissolution favoured by the decrease of the electrostatic force between crystal ions because of high permittivity of the solvents.
Examples	NaCl, CsCl, LiF, Ba SO_4, Na_2SO_4.

Since the electrostatic field of an ion extends in all directions, the process of Na^+ ion attracting a Cl^- ion in the formation of NaCl does not stop at the two ions only. As shown in Fig. 19.39, a positive

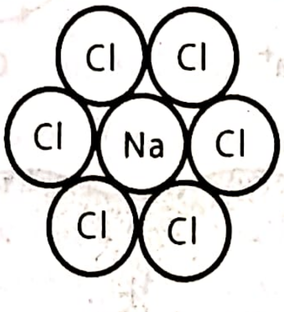


Fig. 19.39

Na ion will be surrounded by negative Cl ions. Likewise, the negative Cl ions will be surrounded by positive Na ions. This leads to the crystalline structure of NaCl belonging to the cubic system having FCC lattice.

(ii) The covalent bond : In covalent bond, atoms are held together by the sharing of electrons. Each atom participating in a covalent bond contributes electrons to the bond. These electrons are shared by all the atoms participating in the bond rather than becoming the exclusive property of one of the atoms as in an ionic bond. Diamond is a crystal whose atoms are linked by covalent bonds. The four outer (valence) electrons in each carbon atom will be shared with those of adjacent carbon atoms giving rise to strong directional bonds. The covalent bond is, therefore, characterized by marked directional properties. For example, the carbon atoms joining each other or other atoms by four bonds make tetrahedral angles with each other. Thus, in a diamond crystal, which is a typical example of the covalent bond, each carbon atom will be at the centre of the tetrahedron formed by the nearest neighbouring atoms. The covalent bond is usually formed by two electrons, one from each atom taking part in the bond. These electrons have anti-parallel spin.

Methane (CH_4) is another example of covalent bonding. Fig. 19.40 illustrates the formation of methane molecule. Carbon has 4 valence

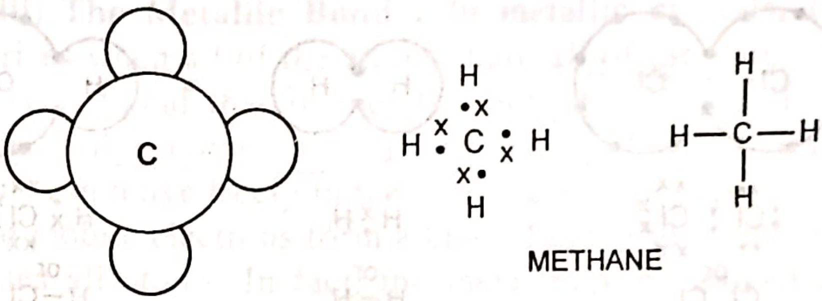


Fig. 19.40

electrons. Each of these electrons forms an electron pair with the single (valence) electrons of the four surrounding hydrogen atoms.

In the formation of covalent bond, a certain degree of stability is achieved as a result of the filling up of the valence shell of the participating atoms due to sharing of the electrons. Consider the case of Cl atoms. As shown in Fig. 19.41, each chlorine atom has 7 valence electrons. Each chlorine atom contributes one electron and the electron

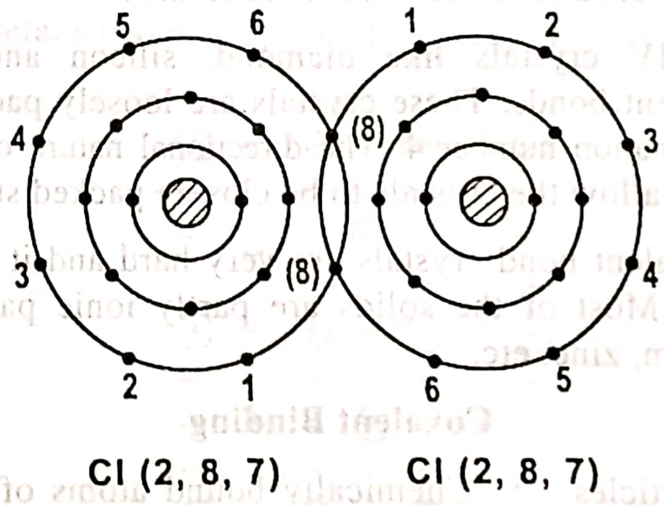


Fig. 19.41

pair is shared by the two atoms so that each has 8 electrons in its valence shell. Thus the chlorine atom achieves argon structure. Similarly, covalent bond is established between two hydrogen atoms by sharing of their electrons resulting in each hydrogen atom acquiring He structure. Covalent bond between hydrogen and chlorine can also be established by sharing of electrons. In this case hydrogen acquires helium structure while chlorine acquires argon structure. Fig. 19.42

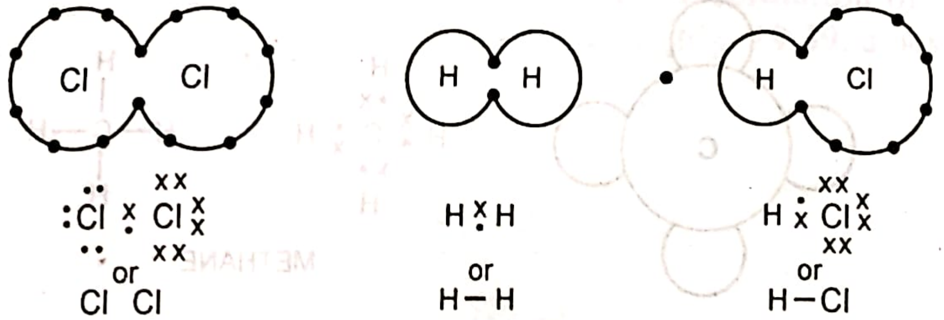


Fig. 19.42

illustrates these formations where the two shared electrons are shown by dots and crosses. It must, however, be emphasized that *attainment of noble gas or octet configuration is not essential in the formation of a covalent bond.* The essential feature of such a bond is the fact that it involves the *pairing of two electrons with opposite spin.* A covalent bond may be either polar or non-polar depending on the fact whether the electron pair is shared unequally or equally between the two atoms. When more than one pair of electrons is shared between two atoms, a double or triple covalent bond is said to be formed.

Group IV crystals like diamond, silicon and germanium possess covalent bonds. These crystals are loosely packed structure with a co-ordination number 4. The directional nature of the covalent bond does not allow the crystals to be closely packed structures.

The covalent bond crystals are very hard and it is difficult to deform them. Most of the solids are partly ionic partly covalent, *e.g.*, germanium, zinc, etc.

Covalent Binding

Constituent particles	:	Chemically bound atoms of one or more kinds.
Binding forces	:	Covalent bond forces are highly directional in nature.
Properties	:	Very hard, have high melting points and insoluble in all ordinary liquids but soluble in non-polar solvents such as benzene and carbon tetrachloride. Poor conductors of heat and electricity and high latent heat of fusion.
Examples	:	Silicon, quartz, diamond, etc.

(iii) The Metallic Bond : In metallic crystals, the metallic bond arises when all of the atoms share all of the valence electrons. It has been found that in metals, each atom loses all its valence electrons and becomes a positively charged ion. These detached electrons can move freely in the interstices of positive ions. Thus the free and mobile electrons form a kind of electron *cloud* or *gas* which permeates all atoms. In fact, the metal may be looked upon as an array of closely-packed ions immersed in a sea of free electrons. The metal is held together by the electrostatic attraction between the positive ions and the negative valence electrons passing between them. The electrons are not bonded directly to an individual atom but they move freely in the sphere of influence of other atoms and are bound to *different atoms at different times* and that too for part of the time.

A metallic bond may be regarded as an *unsaturated* covalent which is due to the electrostatic attraction between the negative electron gas and the positive ion cores. The ion cores consist of the nucleus and the non-valence electrons. Fig. 19.43 illustrates the ion cores of a metal surrounded by mobile electrons.

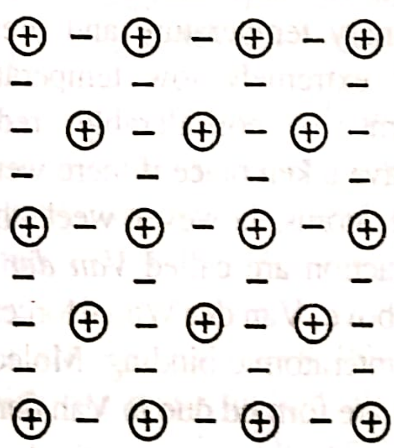


Fig. 19.43

The metallic bond is comparatively weaker than the ionic and covalent bonds which are called *saturated bonds*. Due to the unsaturated nature of the metallic bond, the metals are weak compared to ionic and covalent crystals. But due to the presence of free electrons, metals have high electrical and thermal conductivities. The best example for a metallic crystal is sodium.

Metallic Binding

Constituent particles	:	Positive ions in a sea of free and mobile electrons.
Binding forces	:	Electrostatic force of attraction.
Properties	:	Crystalline structure very soft to very hard possessing metallic luster, ductile and malleable, good conductors of heat and electricity. Moderate to high melting temperatures. Opaque to light.
Examples	:	Metals and their alloys.

(iv) The Molecular Bond : This bonding occurs for those elements or compounds whose electron configuration is such that there is little scope for electron transfer between their atoms. For example, noble gases like helium, the outer or valence shell is complete with two (2s) electrons. Similarly other noble gases like neon and argon have full compliment of electrons in their outer shells. Consequently, in their case, the three primary bonds considered above cannot be formed. Hence atoms of all noble gases have very little attraction for each other and consequently they remain mono-atomic under ordinary temperature and pressure conditions. These gases condense only at extremely low temperature when the thermal agitation of their atoms is considerably reduced. Obviously, the condensation could not have taken place if there were no inter atomic forces of attraction between the atoms, however week they may be. This weak interatomic forces of attraction are called *Van der Waal's forces.* In cases like the ones considered above, Van der Wall's forces are the only one which operate in the process of interatomic binding. Molecular solids of CH_4, Cl_2, I_2, CO_2, C_6H_6 and paraffin are formed due to Van der Wall's interaction.

The Van der Waal's attraction was first explained for electrically neutral gas molecules by Debye. He assumed that neighbouring molecules induce dipole in each other because of their own changing electrical fields. This interaction produces an attractive force which varies inversely as the seventh power of the interatomic separation. Obviously, the force is very weak. The molecules are located at the crystal lattice points and the binding between them are due to Van der Waal's forces. Since the interatomic force in the inert gas solids is very weak, the inert gas solids have very low melting points *e.g.*, the melting point of helium is 0.8K.

Van der Waal's binding

Constituent particles	:	Small molecules or atoms.
Binding forces	:	Van der Waal's attractive forces
Properties	:	Structure crystalline as well as non-crystalline. Soft, poor conductors of heat and electricity, low latent heat of fusion, low melting point and volatile. Soluble in both polar and non-polar liquids. Usually transparent to light.
Examples	:	Solid carbon dioxide, methane, paraffin, sulphur; ice, Cl_2, I_2, etc.

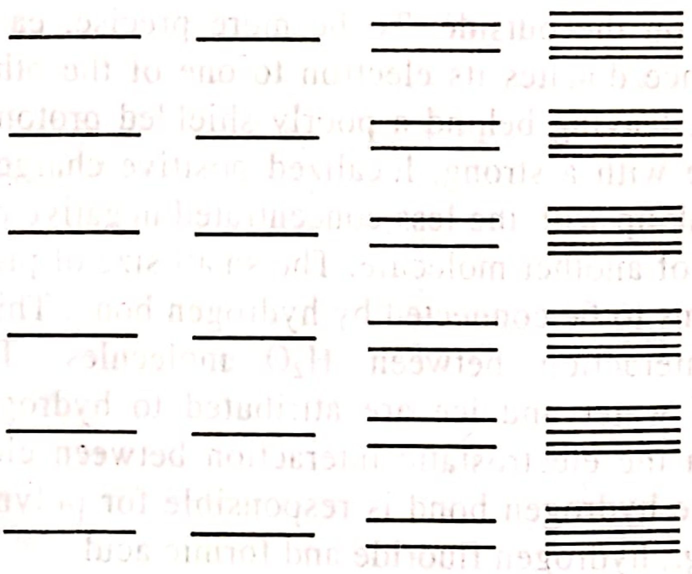
(v) **The Hydrogen Bond :** A hydrogen atom has only one electron and it should therefore form a covalent bond with only one other atom. However, under certain conditions, a hydrogen atom is attracted by rather strong forces towards two atoms. This results in the formation of what is known as *hydrogen bond* between them. The hydrogen bond is believed to be largely ionic in character. However, it is weaker than ionic or covalent bond but stronger than ordinary Van der Waal's bond.

Hydrogen bonds occur in the hydrides of atoms so electro-negative that the molecules may be thought of as having virtually bare protons on the outside. To be more precise, each hydrogen atom in essence donates its electron to one of the other atoms of the molecule, leaving behind a poorly shielded proton. The result is a molecule with a strong, localized positive charge at one end which can link up with the less concentrated negative charge at the opposite end of another molecule. The small size of protons allows only two atoms to be connected by hydrogen bond. This bond is an important interaction between H_2O molecules. The striking properties of water and ice are attributed to hydrogen bonding, together with the electrostatic interaction between electric dipole moments. The hydrogen bond is responsible for polymerisation of molecules *e.g.*, hydrogen fluoride and formic acid.

19.18 Band Theory of Solids

According to theories relating to structure of atom, each electron in an isolated atom must lie in one of a number of permitted orbits. Each permitted orbit has a definite energy associated with it. When isolated, each atom is surrounded by a number of these specific energy levels or permitted orbits. Electrons cannot occupy spaces between these levels. Electrons fill the lowest energy levels first. A specific quantity of energy must be supplied to move an electron to the next higher level. Further, according to Pauli exclusion principle, no two electrons can occupy the same quantum state or energy level. Thus not more than two electrons can occupy any one energy level. Two electrons can occupy the same energy level because their spin (or magnetic quantum number) are in the opposite direction. Thus if we consider the case of a sodium atom, there are 11 electrons. Their distribution among the various energy levels is represented as $1s^2$, $2s^2$, $2p^6$, $3s^1$. According to Pauli exclusion principle, the 10 electrons will fill up the first two shells - the K and L shell. The remaining single electron will have to be in the M shell.

Let us consider that two similar isolated atoms are brought very close to each other. The interaction or coupling between the orbits of their electrons causes each individual energy level to be split up into two slightly different energy levels. When n identical atoms are brought close to each other, each energy level splits up into n slightly different energy levels. Fig. 19.44 illustrates this splitting of energy levels.



(a) (b) (c) (d)

Fig. 19.44

Now in a crystal, this number n is very large. In a sodium chloride crystal $n = 10^{23}$. Hence when these n number of atoms are brought close to each other, each discrete energy level splits into n numbers of very closely-spaced levels of slightly different energy. This set of closely spaced energy levels of slightly different energies is called an energy band. The individual energies within the band are so close together that, for many purposes, the energy band may be considered to be continuous. Fig. 19.45 illustrates the splitting of K, L and M levels as the distance between different atoms is reduced. At

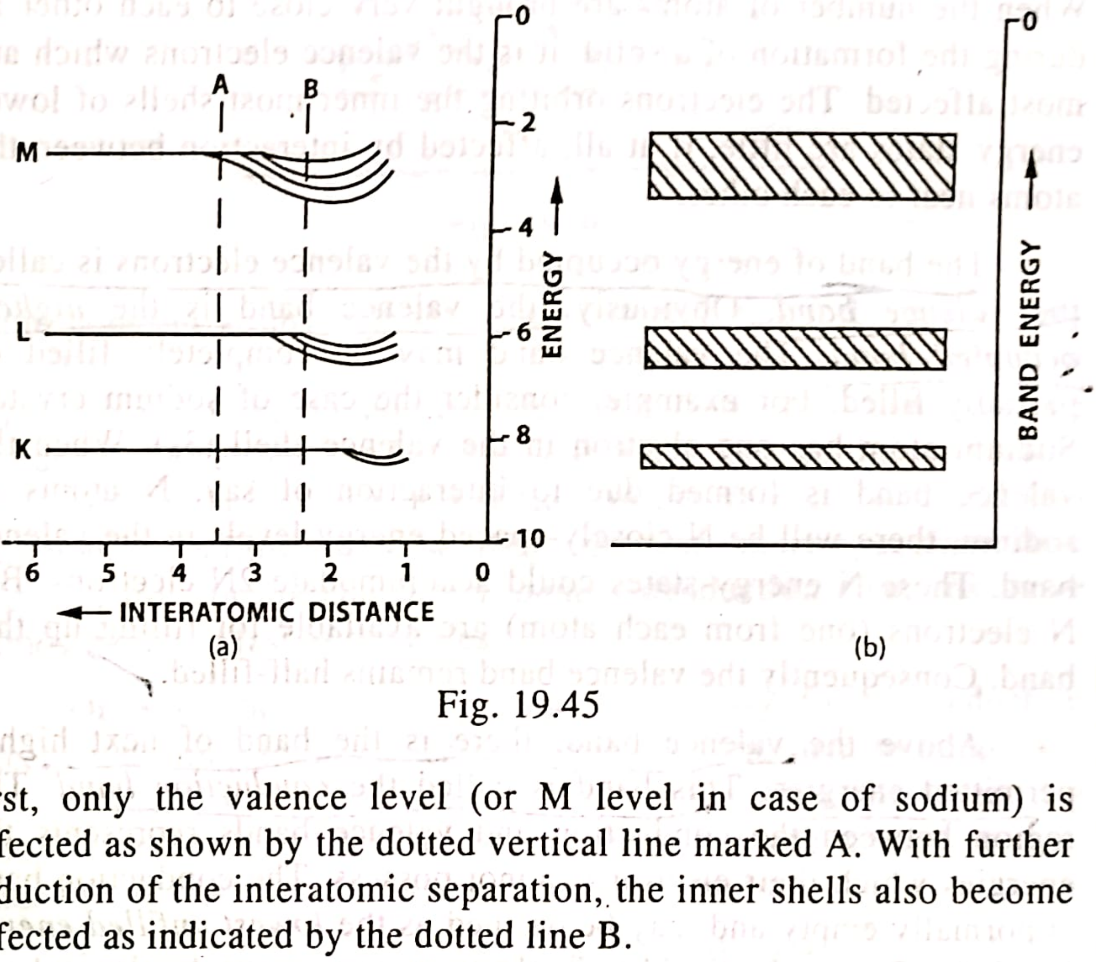


Fig. 19.45

first, only the valence level (or M level in case of sodium) is affected as shown by the dotted vertical line marked A. With further reduction of the interatomic separation, the inner shells also become affected as indicated by the dotted line B.

The energy bands in a crystal correspond to energy levels in an atom and an electron in a crystal can only possess an energy that falls within one of these bands. Depending on the structure of the crystal, the various energy bands of a crystal may or may not overlap. If they do not overlap, they are separated by regions which have no allowed energy levels *i.e.*, the region represent energy values which the electrons in the crystal can not have. These regions are known as *forbidden bands* or *energy gaps*.

If there are N atoms in a solid, there will be N allowed quantum states or energy levels in each band. A maximum of two electrons (spin up and spin down) can occupy each quantum state. Thus each energy band can be occupied by 2N electrons.

Valence and Conduction Bands

The outermost electrons of an atom *i.e.*, those in the shell furthest from the nucleus are called *valence electrons*. These electrons have the highest energy (algebraically not numerically). When the number of atoms are brought very close to each other as during the formation of a solid, it is the valence electrons which are most affected. The electrons orbiting the inner most shells of lower energy states are little, if at all, affected by interaction between the atoms near to each other.

The band of energy occupied by the valence electrons is called the *valence band*. Obviously, the valence band is the *highest occupied band*. The valence band may be completely filled or partially filled. For example, consider the case of sodium crystal. Sodium atom has one electron in the valence shell (3s). When the valence band is formed due to interaction of say, N atoms of sodium, there will be N closely-spaced energy levels in the valence band. These N energy-states could accommodate 2N electrons. But N electrons (one from each atom) are available for filling up this band. Consequently the valence band remains half-filled.

Above the valence band, there is the band of next higher permitted energies. This band is called the *conduction band*. The region between the conduction and valence bands represents the energies which their electrons cannot possess. The conduction band is normally empty and may be defined as the *lowest unfilled energy band*. In the conduction band, electrons can move freely and are generally called *conduction electrons*. The gap between the conduction and valence bands is called the *forbidden energy gap*.

As discussed earlier, the covalent forces of the crystal lattice have their source in the valence band. If by some means, the valence electron happens to absorb enough energy, it can jump across the forbidden energy gap and enter into the conduction band. When an electron is ejected from the conduction band, a covalent bond is

broken and a positively charged hole is left behind. This hole can travel to an adjacent atom by acquiring an electron from that atom which involves breaking of an existing covalent band. However, when the hole is filled up, the covalent band is re-established. As illustrated in Fig. 19.46, the holes are filled by electrons which move

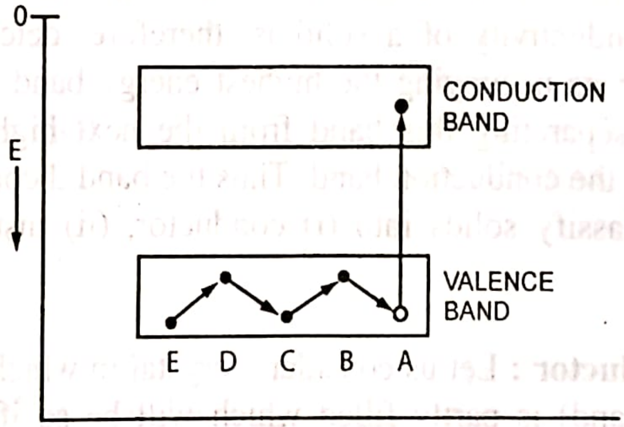


Fig. 19.46

from adjacent atoms without passing though the forbidden energy gap. This is another way of saying that conditions in the conduction band have nothing to do with the hole flow. To summarize, (i) conduction electrons are found in and freely flow in the *conduction* band and (ii) holes exist in and flow in the valence band and (iii) although holes flow with ease, they *experience more opposition than electron flow in the conduction band* - conduction electrons move almost twice as fast as the holes.

Once in the conduction band, it is easy for a conduction electron to jump to an adjacent conduction band than to jump back to the valence band from where it had come earlier. However, if a conduction electron happens to radiate too much energy, it will suddenly reappear in the valence band once again.

At the absolute zero of temperature all the energy levels below a certain level will be filled with electrons and all levels above this will be empty. The energy level that divides the filled and empty levels is referred to as *Fermi level.*

Conductors, insulators and semi-conductors :

The electrical conduction properties of different elements and compounds can be explained in terms of band theory. The bands which

are formed by the electrons of the completely-filled electronic shells of an atom are *exactly full* as there are even numbers of electrons in these shells. An applied electric field cannot produce any change in the state of the electrons in a completely filled band. The electrons of the completely filled band cannot, therefore, contribute to electrical conductivity. The bands formed from the inner electronic shells are narrow and do not overlap. The conductivity of a solid is, therefore, determined by the number of electrons occupying the highest energy band (valence band) and the region separating this band from the next higher unoccupied energy band *i.e.*, the conduction band. Thus the band theory of solids can be applied to classify solids into (i) conductor, (ii) insulator and (iii) semi-conductor.

(i) **Conductor :** Let us consider a crystal in which the outermost band (valence band) is partly filled which will be so if the outermost shell has odd number of electrons. As the band is partly vacant, it has vacant sites into which the electron can move when it acquires an additional energy, say, when an electric field is applied. Let us take the case of sodium (Z=11). In an isolated atom these eleven electrons have the configuration of $1s^2$, $2s^2$ $2p^6$, $3s^1$. The ten electrons form closed shells, hence they are expected to form vary narrow bands in the solid. These bands will be completely filled. The single outermost electron per atom will only fill half of the next band *i.e.*, the valence band. Fig. 19.47 illustrates the energy bands in sodium. Now when an electric field is applied across a solid sodium, the electrons in the valence band easily acquire the required additional energy to move into the *higher unoccupied energy levels within the same band without crossing any energy gap*. The additional energy is in the form of kinetic energy, and the moving electrons constitute an electric current. Sodium metal is therefore a good conductor of electricity.

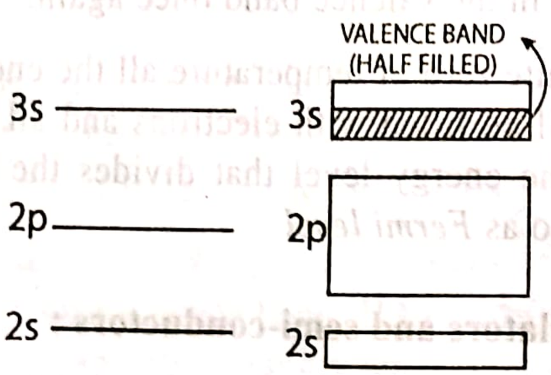


Fig. 19.47

All sodium like alkali metals which have one valence electron per atom is a metal and a good conductor. The alkaline earth metals *e.g.*, Magnesium (Z=12) have electronic configuration of $1s^2$, $2s^2$ $2p^6$, $3s^2$. Magnesium therefore has two valence electrons per atom. The valence band is therefore completely filled and magnesium should be a non-conductor (insulator) of electricity. But the full valence band overlap the conduction band as illustrated in Fig. 19.48. Here the forbidden energy gap is zero. This renders magnesium a metal but not very good metal. Thus magnesium and other alkaline earth metals is conductor but not very good conductors.

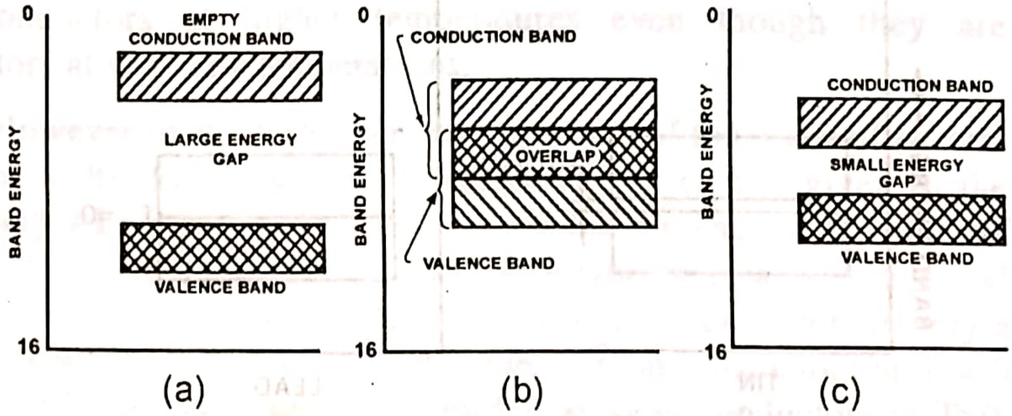


Fig. 19.48

(ii) Insulators : Fig. 19.49 illustrates the energy band diagram of diamond – a form of carbon. The number of electrons in the outermost shell of carbon is even and hence the valence band is empty at absolute zero of temperature. In the case of diamond, the energy gap between the two bands is very large – nearly 6 eV as shown in the figure. The only way the electrons conduct electric current is by transition across the forbidden gap into the conduction band. This means that at least 6 eV of additional energy must be provided to an electron in a diamond crystal if it has to make this transition. This can be brought about *by applying a very strong field or by thermal excitation*. However, it has been found that an electric field of many millions of volts per metre is necessary to make the

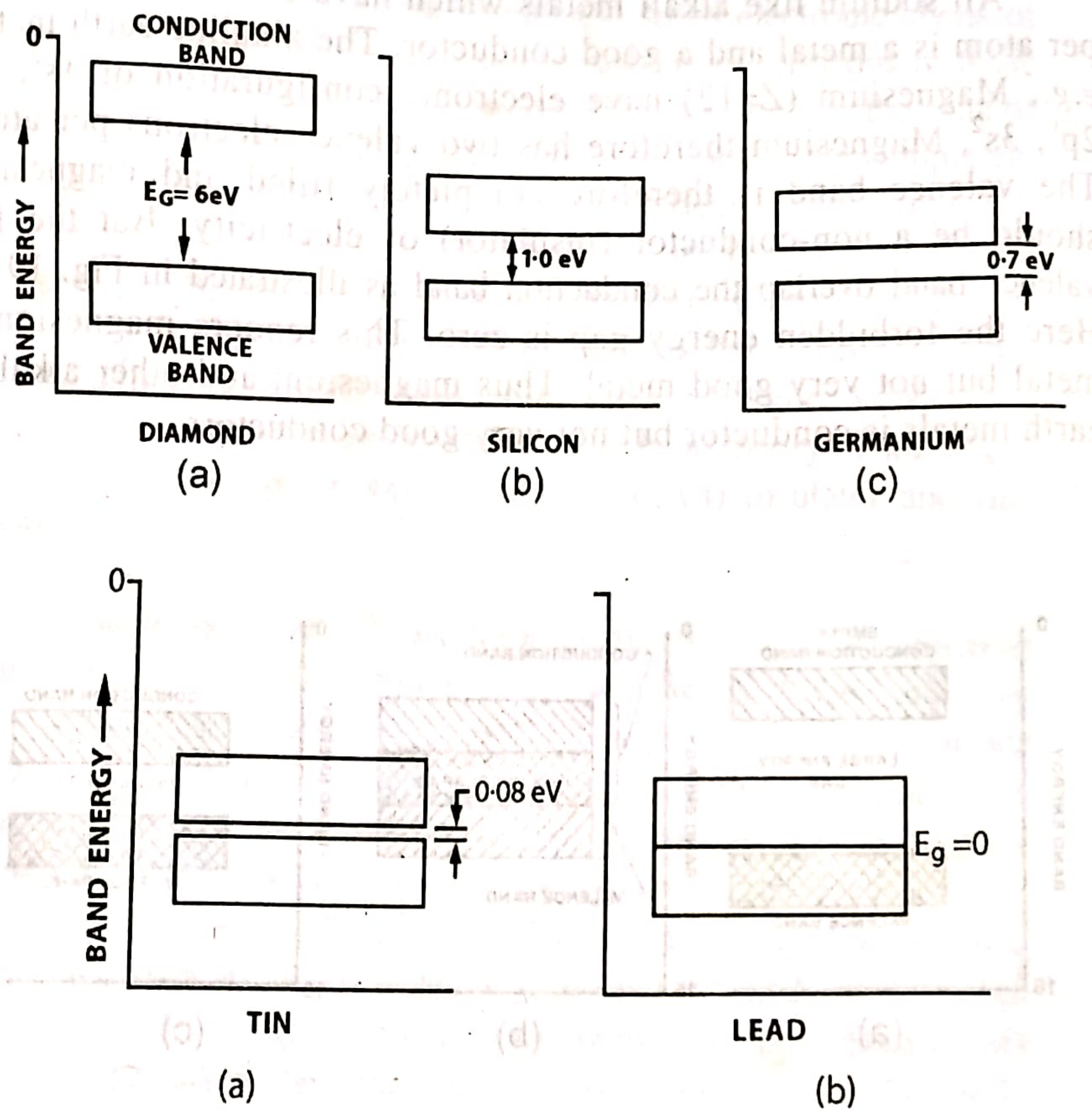


Fig. 19.49

electrons jump across the forbidden gap. This is normally not possible. Also at normal room temperature, the number of electrons that can be thermally excited across the gap in diamond is very small. Thus the necessary energy cannot be provided either by electric field or temperature. Therefore, there will be no electric current and as such diamond is regarded to be an insulator.

(iii) Semi-conductor : Silicon and germanium have crystal structures resembling that of diamond. As in diamond, an energy gap separates the top of the filled valence energy band from a vacant higher band (conduction band). According to definition, silicon and germanium should be insulators, which the crystals are at their lowest energy states. But the width of the forbidden gap in silicon is 1.1 eV, it

is even less (0.7 eV) in case of germanium. Now let us consider the case when T>0. At temperatures different from zero some electrons from the upper filled band can be excited into the next empty (conduction) band and conduction becomes possible. At ordinary room temperatures a small proportion of electrons in the valence band of silicon and germanium will have sufficient kinetic energy due to thermal excitation to cross the narrow energy gap and enter into the conduction band above it. These electrons are sufficient to permit a limited amount of current to flow when an electric field is applied. Thus silicon as well as germanium have electrical resistivity between that of conductors and insulators. Hence these are termed *semiconductors*. Solids, such as silicon and germanium are conductors at higher temperatures even though they are insulators at very low temperatures.

However, if the forbidden energy gap is of the order of several electron volts, as in the case of diamond, the energy given to the electron $\simeq kT$ is not sufficient to overcome the energy gap and the solid remains an insulator. It should, however, be clearly understood that the distinction between insulators and semi-conductors is only a quantitative one. In fact all semi-conductors are insulators at T = 0 whereas all insulators may be regarded as semi-conductors at T>0. This explains why the conductivity of semi-conductors increases with temperatures. As temperature increases, more and more electrons go from the valence band to the conduction band thereby increasing the conductivity. However, in a metallic conductor increase in temperature also agitates the electrons, but these electrons are already in the conduction band and free for conduction. Thus the effect of the increased agitation of the free electrons is to impede the progress of the electron flow when an electromotive force is applied to the conductor, due to more collisions taking place. Thus, an increase in resistance occurs with an increase in temperature in most metals, *i.e.*, the conductivity decreases with increase in temperature.

Atomic Binding in Semi-conductor

Semiconductors like germanium and silicon, have crystalline structure *i.e.*, their atoms are arranged in an ordered array known as crystal lattice. Both germanium and silicon are tetravalent *i.e.*, each

has four valence electrons in the outermost shell. A single germanium atom has actually 32 electrons, 28 of which are tightly bound to the nucleus. The remaining four electrons revolve in the outer most shell and are called valence electrons. These valence electrons are responsible for the electrical properties of germanium. Now the neighbouring atoms in both germanium and silicon crystals form covalent bonds by sharing four electrons with each other so as to achieve the inert gas structure (*i.e.*, 8 electrons in the outermost orbit). Fig. 19.50 illustrates the two-dimensional view of the germanium crystal lattice. The circles represent cores of the atom

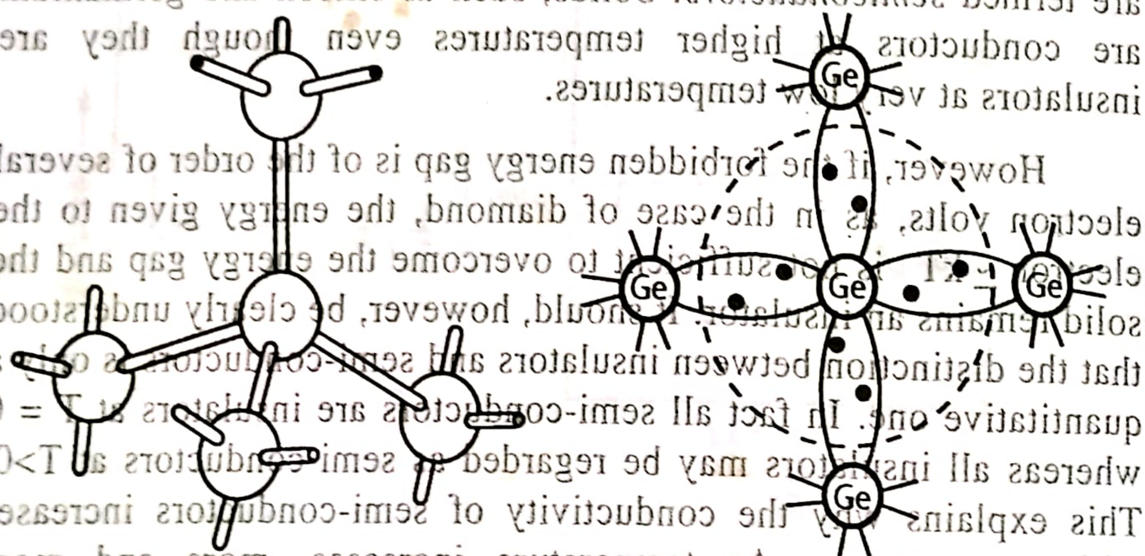


Fig. 50

consisting of the nuclei and 28 electrons. Each pair of lines represent a covalent bond with the dots representing the valence electrons. As can be seen, each atom has 8 electrons under its influence. Fig. 19.50 illustrates a 3-dimensional view of the germanium crystal where each atom is surrounded symmetrically by four other atoms forming a tetrahedral crystal. Each atom shares a valence electron with each of its four neighbours, thereby forming a stable structure.

Intrinsic semiconductor

Germanium, silicon and other semiconductors in their extremely pure form are called *intrinsic* semiconductors. At absolute zero of temperature, the valence band of a pure (intrinsic)

semiconductor is completely full and the conduction band completely empty. The covalent bond formed by the neighbouring atoms of an intrinsic semiconductor must be broken to provide electrons for conduction. One way to rupture the covalent bond and set the electron free is to increase the crystal temperature above 0°K. Even at ordinary temperature, an appreciable number of electrons will be excited to the conduction band. Thus by thermal excitation electrons can be freed from the valence band and transferred to the conduction band across the forbidden energy gap. The electrons thus transferred to the conduction band can conduct electricity. Moreover, electrons leaving the valence band leave behind some vacancies or *holes* in the valence band. This hole is filled up by a nearby electron which in turn leave one more hole behind. Thus electron continue to jump from hole to hole in a valence band in the same direction as a free electron in the conduction band. The effect is that of a hole or a positive charge (equal to that of an electron) moving in the opposite direction. This is called the *hole current* as distinct from the normal electronic or conduction current. When a potential difference is applied between the two faces of a semiconductor, electrons drift towards +ve end and holes towards the -ve end. Hence, semiconductor current is the sum of electron and hole currents flowing in the opposite direction.

It must be clearly understood that holes do not leave or enter the semiconductor because there are no holes in the connecting wires, but the hole current is maintained by electrons entering and leaving the semiconductor.

It can be proved that in an intrinsic semiconductor Fermi energy level lies in the middle of the energy gap *i.e.*, midway between the conduction and valence bands.

Extrinsic semiconductor

Silicon, germanium and other semiconductors in their extremely pure form are called intrinsic semiconductors. In an intrinsic semiconductor there will be as many holes as there are free electrons. By introducing tiny amount of foreign or impurity elements in an intrinsic semiconductor, imperfections are produced in the crystal lattice, and the number of holes (or electrons) can be made greater than the number of free electrons (or holes) depending

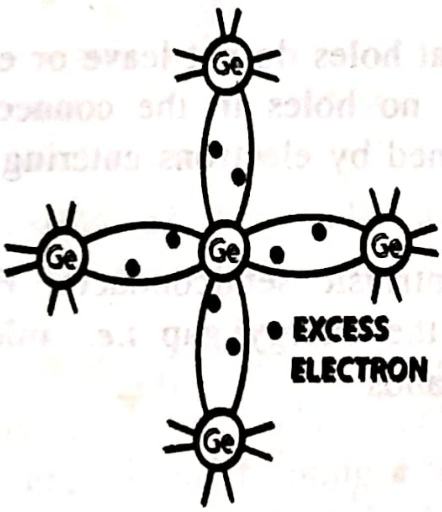
on the nature of the impurity added. This drastically alters the electrical conductivity of the semiconductors. Since the electrical conductivity of the semiconductors now depend strongly upon the impurity content, the semiconductors are now called *extrinsic semiconductor*. The process of adding impurities is called *doping* and the semiconductor is said to have been *doped*.

Depending upon the nature of the doping material, extrinsic semiconductors can be divided into two classes:

(i) N-type semiconductor and (ii) P-type semiconductor.

N-type semiconductor:

This type of semiconductor is obtained when a pentavalent material like phosphorous, arsenic or antimony is added to pure germanium (or silicon) crystal. Suppose a pentavalent atom such as antimony is substituted for a germanium atom in the crystal. As shown in Fig. 19.51, four of the impurity atom's electron play the same role as the four valence electrons of the replaced germanium atom.



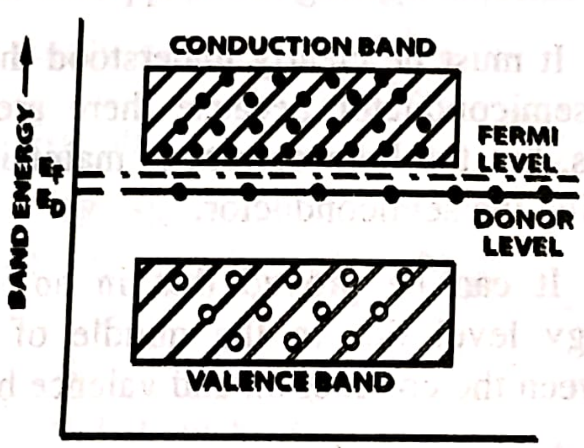


Fig. 19.51

In other words, each antimony atom forms covalent bonds with the surrounding four germanium atoms with the help of four of its five electrons. The fifth electron can find no electron to pair with and

becomes superfluous and is losely bound to the antimony atom. In terms of energy level, the fifth antimony electrons has an energy level (called donor level) just below the conduction band. Usually, the donor level is 0.01 eV below conduction band for germanium and 0.054 eV for silicon. Thus the fifth electron from the antimony atom can be easily detached from the parent atom by thermal excitation and it moves freely into the conduction band. One electron is therefore, present in the conduction band for each impurity atom in the crystal *without creating a positive hole*. Hence, there will be a large increase in the number of free electrons in the crystal. Since antimony has been responsible for donating electrons to the crystal, it is called a *donor* impurity and the germanium crystal is said to have been doped. Germanium has now become an extrinsic semiconductor.

As in an intrinsic crystal, thermal excitation still causes electron hole pairs, but, because of the abundance of the donated electrons, combination of holes and electrons occurs rapidly, so that there are very few holes in the doped crystal (even less than in pure germanium). Because the free electron density is predominant (over holes), the crystal is called *donor* or *emitter*, and the crystal is called a *N-type* crystal (N for negative) or semiconductor. Since the concentration of electrons in the conduction band far exceeds the concentration of holes in the valence band, the Fermi level shifts upwards to the bottom of the conduction band as shown in Fig. 19.51.

It can be seen from the above description that current flow through a N-type semiconductor is due primarily to the donated electrons in the conduction band, so that the electrons are called the *majority current carriers*. The holes are termed *minority current carriers*. The holes are termed minority current carriers, since the current flow in the valence band is very small. The conductivity of pure germanium (or silicon) can be drastically increased (approximately 1000 times) by the addition of small but controlled amount of impurity material (1 atom per 10 million atoms of germanium) into the crystal lattice.

It may be noted that by giving away its one valence electron, the donor atom becomes positively charged. But it cannot take part in conduction because it is firmly fixed or tied into the crystal lattice. But because the donor impurity brings in as much negative charge (by way of electrons) as positive charge (by way of protons), the total charge of the semiconductor does not change.

P-type semiconductor

If, instead of a pentavanlent material, a minute quantity of a trivalent material such as boron, aluminium, gallium or indium, is introduced (in the ratio of 1 atom for every 10^6 atoms of germanium), the opposing doping effect will be obtained. In this case, the three valence electrons of the impurity atom, say boron, form covalent bonds with three atoms of the surrounding germanium atoms but one bond is left incomplete. This deficiency of an electron gives rise to an electron hole at the location of the impurity which breaks up the crystal lattice. An electron from the adjacent germanium atom may migrate to the vacancy at the site of the impurity (Fig. 19.52). The foreign atom is said to have accepted a valence electron and the crystal is called an *acceptor* or *collector*. When the impurity atom accepts an electron for bonding purpose it

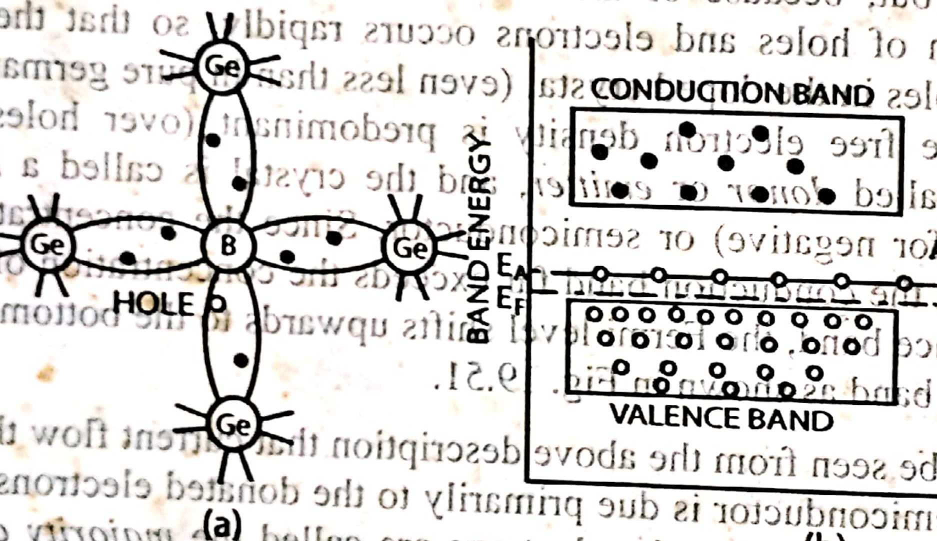


Fig. 19.52

acquires a bound negative charge. The hole that was created in the adjoining germanium atom is available to be filled by another electron from another atom. Hence the hole moves through the crystal. At room temperature, thermal excitation causes electron-hole pairs to be formed as in an intrinsic crystal. But because of greatly increased number of holes (there are as many additional holes as there are impurity atoms), there are very few electrons (even less than in pure germanium). *Since the holes predominate they are called the majority current carriers while the free electrons are the minority carriers.* The hole behaves as if it is a positive charge and consequently the crystal is called a *P-type crystal.*

From the above description it is obvious that in a P-type material, most current flows by means of electrons filling holes in the valence band. Thus the mechanism of current flow in a P-type material is quite different from that in a N-type material where the current flows by the motion of electrons in the conduction energy band. This produces an extremely small current in a P-type material. As can be expected Fermi level shifts nearer to the valence band (Fig. 19.52), since the number of holes in the valence band is greater than the number of electrons in the conduction band. The acceptor level lies immediately above the Fermi level. Since the conduction is by means of hole movement at the top of the valence band, the acceptor level readily accepts the electrons from the valence band.

Again it may be noted that although a P-type semiconductor has excess of holes for conduction purposes, the material as a whole is electrically neutral as explained in the case of a N-type material.

EXERCISE

1. What is a crystal? Define a unit cell. What are the lattice parameters of a unit cell?
2. What are Bravais lattices? Describe the different types of crystal systems along with their characteristics.
3. What is space lattice? Describe briefly the seven systems of crystals.
4. Explain the term symmetry in the structure of crystals. Describe the different symmetry elements.
5. What is co-ordination number? What are the co-ordination number for a (i) simple cubic, (ii) body centred cubic and (iii) face centred cubic structure of crystal.
6. What do you mean by dimensions of unit cell? Calculate the dimensions of unit cell for a body-centred cubit lattice and a face-cintred cubic lattice.

7. Show that for a b.c.c. and f.c.c. crystal structure, the lattice constants are given by

$$a_{bcc} = \frac{4r}{\sqrt{3}} \quad \text{and} \quad a_{fcc} = \frac{4r}{\sqrt{2}}$$

where r is the atomic radius.

8. What is packing fraction? Derive expressions for the pacing fraction of a s.c, b.c.c. and f.c.c. crystal structure.

9. Briefly describe the important plane systems in a cubic crystal.

10. What are Miller indices? Show that in a crystal of cubic structure, the distance between the planes with Miller indices h, k, l is equal to do $= \frac{a}{\sqrt{h^2+k^2+l^2}}$ where a is the lattice parameter.

11. Draw a schematic diagram of the unit cell for the simple cubic, body-centred cubic and face-centred cubic lattices.

12. Briefly discuss the various defects in a crystal.

13. Distinguish between Frenkel and Schottky defects.

14. Briefly describe the different bonds in solids.

15. Explain the formation the different energy bands in a solid and hence distinguish between an insulator, a semiconductor and a conductor.

16. Explain with the help of energy band diagrams, the difference between an intrinsic and an extrinsic semiconductor giving examples.

17. Explain the following:

(i) Ionic binding (ii) covalent binding, (iii) Metallic binding and (iv) Van der Wall's binding.

18. What is the difference between N-type and P-type semiconductors? Explain with the help of diagrams.

19. What is an intrinsic semiconductor? Give one example.

Explain with the help of a diagram how you can convert an intrinsic semiconductor into an n-type or p-type extrinsic semiconductor.

20. Discuss the function of an impurity in (i) N-type and (ii) P-type semiconductor.

by the same author

Practical Physic

for degree students

(BSC pass honours and Engineering students)

Advanced Practical Physics

Physics for Engineers Part -1

Physics for Engineers Part -2

by the same Publishers

Engineering Materials - M.A.Aziz

A Test Book of Surveying - M. A. Aziz & M. Shahjahan

A Test Book of Water Supply Engineering - M.A.Aziz